UNIL | Université de Lausanne
Faculté de biologie et de médecine

Département d'Écologie et d'Évolution

Using environmental niche modeling to understand biological invasions in a changing world

Thèse de doctorat ès sciences de la vie (PhD)

présentée à la

Faculté de biologie et de médecine de l'Université de Lausanne

par

Blaise PETITPIERRE

Master de l'Université de Lausanne

Jury

Prof. Sophie Martin, Présidente
Prof. Antoine Guisan, Directeur de thèse
Dr. Jake Alexander, Expert
Dr. Wilfried Thuiller, Expert

Lausanne 2013

UNIL | Université de Lausanne
Faculté de biologie
et de médecine

Ecole Doctorale

Doctorat ès sciences de la vie

Imprimatur

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

Présidente	Madame	Prof. Sophie **Martin**
Directeur de thèse	Monsieur	Prof. Antoine **Guisan**
Experts	Monsieur	Dr Jake **Alexander**
	Monsieur	Dr Wilfried **Thuiller**

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

Monsieur Blaise Petitpierre

Master of Science de l'Université de Lausanne

intitulée

Using environmental niche modeling to understand biological invasions in a changing world

Lausanne, le 31 janvier 2014

pour La Doyenne
de la Faculté de Biologie et de Médecine

Prof. Sophie Martin

RÉSUMÉ

L'activité humaine affecte particulièrement la biodiversité, qui décline à une vitesse préoccupante. Parmi les facteurs réduisant la biodiversité, on trouve les espèces envahissantes. Symptomatiques d'un monde globalisé où l'échange se fait à l'échelle de la planète, certaines espèces, animales ou végétales, sont introduites, volontairement ou accidentellement par l'activité humaine (par exemple lors des échanges commerciaux ou par les voyageurs). Ainsi, ces espèces atteignent des régions qu'elles n'auraient jamais pu coloniser naturellement. Une fois introduites, l'absence de compétiteur peut les rendre particulièrement nuisibles. Ces nuisances sont plus ou moins directes, allant de problèmes sanitaires (p. ex. les piqûres très aigües des fourmis de feu, originaires d'Amérique du Sud et colonisant à une vitesse fulgurante les USA, l'Australie ou la Chine) à des nuisances sur la biodiversité (p. ex. les ravages de la perche du Nil sur la diversité unique des poissons Cichlidés du Lac Victoria). Il est donc important de pouvoir prévenir de telles introductions. De plus, pour le biologiste, ces espèces représentent une rare occasion de pouvoir comprendre les mécanismes évolutifs et écologiques qui expliquent le succès des envahissantes dans un monde où les équilibres sont bouleversés.

Les modèles de niche environnementale sont un outil particulièrement utile dans le cadre de cette problématique. En reliant des observations d'espèces aux conditions environnementales où elles se trouvent, ils peuvent prédire la distribution potentielle des envahissantes, permettant d'anticiper et de mieux limiter leur impact. Toutefois, ils reposent sur des hypothèses pas évidentes à démontrer. L'une d'entre elle étant que la niche d'une espèce reste constante dans le temps, et dans l'espace. Le premier objectif de mon travail est de comparer si la niche d'une espèce envahissante diffère entre sa distribution d'origine native et celle d'origine introduite. En étudiant 50 espèces de plantes et 168 espèces de Mammifères, je démontre que c'est le cas et que par corolaire, il est possible de prédire leurs distributions.

La deuxième partie de mon travail consiste à comprendre quelles seront les interactions entre le changement climatiques et les envahissantes, afin d'estimer leur impact sous un climat réchauffé. En étudiant la distribution de 49 espèces de plantes envahissantes, je démontre que les montagnes, régions relativement préservée par ce problème, deviendront bien plus exposées aux risques d'invasions biologiques. J'expose aussi comment les interactions entre l'activité humaine, le réchauffement climatique et les espèces envahissantes menacent la vigne sauvage en Europe et propose des zones géographiques particulièrement adaptée pour sa conservation.

Enfin, à une échelle beaucoup plus locale, je montre qu'il est possible d'utiliser ces modèles de niches le long d'une rivière à une échelle extrêmement fine (1 mètre), potentiellement utile pour rationnaliser des mesures de conservations sur le terrain.

SUMMARY

Biodiversity is significantly negatively affected by human activity. Invasive species are one of the most important factors causing biodiversity's decline. Intimately linked to the era of global trade, some plant or animal species can be accidentally or casually introduced with human activity (e.g. trade or travel). In this way, these species reach areas they could never reach through natural dispersal. Once naturalized, the lack of competitors can make these species highly noxious. Their effect is more or less direct, from sanitary problems (e.g. the harmful sting of Fire Ants, originating from South America and now spreading throughout USA, China and Australia) or can affect biodiversity (e.g. the Nile perch, devastating the one of the richest hotspot of Cichlid fishes diversity in Lake Victoria). It is thus important to prevent such harmful introductions. Moreover, invasive species represent for biologists one of the rare occasions to understand the evolutionary and ecological mechanisms behind the success of invaders in a world where natural equilibrium is already disturbed.

Environmental niche models are particularly useful to tackle this problematic. By relating species observation to the environmental conditions where they occur, they can predict the potential distribution of invasive species, allowing a better anticipation and thus limiting their impact. However, they rely on strong assumption, one of the most important being that the modeled niche remains constant through space and time. The first aim of my thesis is to quantify the difference between the native and the invaded niche. By investigating 50 plant and 168 mammal species, I show that the niche is at least partially conserved, supporting for reliable predictions of invasive' s potential distributions.

The second aim of my thesis is to understand the possible interactions between climate change and invasive species, such as to assess their impact under a warmer climate. By studying 49 invasive plant species, I show that mountain areas, which were relatively preserved, will become more suitable for biological invasions. Additionally, I show how interactions between human activity, global warming and invasive species are threatening the wild grapevine in Europe and propose geographical areas particularly adapted for conservation measures.

Finally, at a much finer scale where conservation plannings ultimately take place, I show that it is possible to model the niche at very high resolution (1 meter) in an alluvial area allowing better prioritizations for conservation.

REMERCIEMENTS

Un grand merci à mon superviseur Antoine Guisan qui m'a proposé et incité à faire cette thèse. Sa grande confiance m'a permis de bénéficier d'une liberté de créativité exceptionnelle lors de ce doctorat.

Merci aux membres de mon comité, Wilfried Thuiller, Jake Alexander et Sophie Martin. Je ai été très honoré d'avoir pu disserter avec vous sur un ton aussi constructif. Je tiens aussi remercier aussi mes deux experts présents lors de mon examen de demi-thèse, Michael Nobis et Heinz Müller-Schärer qui m'ont dispensé de précieux conseils. Je suis aussi très reconnaissant à Caroline Beto-Colliard qui m'a apporté une précieuse assitance administrative à la fin de mon doctorat, ainsi qu'à Béatrice Regamey du centre de reprographie de l'UNIL qui me permit d'imprimer le premier jet de cette thèse dans des conditions pas faciles.

Lors de ma thèse, je n'ai que rarement eu l'occasion de réfléchir et d'observer mon sujet *in situ*, sur le terrain, alors que c'est une composante fondamentale pour mieux comprendre les mécanismes régissant l'écologie des espèces. Heureusement, grâce à l'assistanat effectué pour les cours et les stages de Pascal Vittoz, j'ai pu compenser (un peu) cette lacune. Je lui suis très reconnaissant pour les édifiantes excursions passées ensemble et son regard aigu sur l'écologie des plantes. De plus, grâce à la confiance qu'il m'a faite lorsqu'il m'a proposé de le remplacer pour son cours, j'ai découvert un plaisir insoupçonné pour l'enseignement. Je suis aussi très reconnaissant envers les étudiants dont j'ai supervisé les travaux de Master, ou de Bachelor, qui m'ont fait sortir hors du bureau pour aller attaquer la problématique des plantes envahissantes : Michael Berthoud, Eros Genilini, Patrice Descombes et Mila Pajkovic.

Par corolaire, ayant passé la majorité de cette thèse devant mon ordinateur au bureau, l'environnement social y était très important. Mes collègues directs du groupe y ont largement contribué : Loïc Pellissier, Anne Dubuis, Jean-Nicolas Pradervand, Julien Pottier, Carmen Cianfrani, Wim Hordijk et Eric Pinto. Un merci tout particulier à Olivier Broennimann, l'autre membre de la « Team Invasive » dans le groupe Ecospat. C'est en partie son travail de thèse que j'ai repris au bond et cette filiation s'est transformée en complicité dépassant largement le cadre professionnel.

Enfin, merci à tous mes proches avec lesquels j'ai passé mes meilleurs moments ces 5 dernières années et qui m'ont permis de tenir dans les moments plus difficiles : mes parents Ariane et José, mon frère Aurèle mais aussi Axelle, François, Diane, Marc, Sylvie, Cyrian, Florent, Kevin et Nadia. Je termine par une pensée pour mon grand-papa Fernand et sa joie de vivre contagieuse, qui n'a pas pu voir l'aboutissement de ce travail mais qui aurait été très heureux d'assister à ma soutenance.

Table of contents

INTRODUCTION

SETTING THE ANTHROPOCENE

Human activities have significantly impacted landscapes and environments, at least at a regional scale, even before the rise of agriculture 10'000 years ago (Pyne, 1995, Steffen *et al.*, 2007). Prehistoric hunters-gatherer populations may be responsible of several waves of extinctions in North America (Martin, 1973), Australia (Roberts *et al.*, 2001) or New Zealand (Holdaway and Jacomb, 2000). Since the appearance of the practice of agriculture, human societies' impact on earth's environment has been considered a significant component of Earth System, i.e. they play a major role in the physical, chemical and biological global-scale cycles and energy fluxes that provide the life-support system for life at the surface of the planet (Kump *et al.*, 2010, Steffen *et al.*, 2007). Preliminary deforestation around 8000 years ago and irrigation of rice 5000 years ago may have counteracted the advent of a potential glacial era through release of carbon dioxide and methane (Steffen *et al.*, 2007). However, the amounts are relatively small compared to the emission of the last 300 years and this hypothesis is vigorously debated (Augustin *et al.*, 2004, Broecker and Stocker, 2006). Nevertheless, it is clear that human societies have started altering Earth System several thousands years ago, but it is only recently that human activities and their impact exploded abruptly and exponentially (Fig. 1), generating such a deep, massive and permanent modification in the Earth System that led to the consideration by some authors that we have entered a new geological era: we left the Holocene (started about 10'000 years ago) to enter the Anthropocene (Crutzen, 2002).

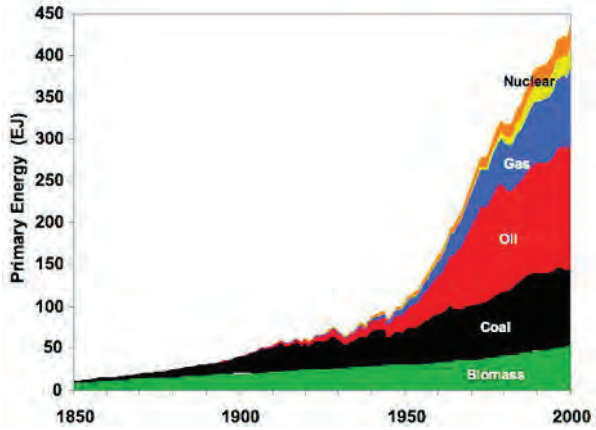


Figure 1. Amount of energy used by mankind and decomposition of the different types of fuels used. From Steffen *et al.* (2007)

The first significant step of this transition was the huge expansion of the use fossil fuels necessary for the industrialization, which started around the 1700s in England and the Low Countries (Steffen *et al.*, 2007). This sudden access to this enormous amount of energy allowed humankind to multiply its demography by a factor six, the global energy consumption by 40 and the global economy by 50 (McNeill, 2000, Steffen *et al.*, 2007) between 1800 and 2000. The consequences of such a human boom have been dramatic for the Earth System, in particular for the component of its biodiversity.

THE SIXTH EXTINCTION

Biodiversity, commonly measured as the number of species in a given area, can be characterized at any moment in time by the difference between speciation (the appearance of new species) and extinction (the disappearance of species). Fluctuation of these rates can provoke increases or decreases in biodiversity. Since the beginning of the Phanerozoic 500 million years ago (i.e. the geologic era when the animal phyla emerged), it is recognized that biodiversity has grown globally (Fig. 2). However, this apparent global increase has been interrupted by abrupt mass extinctions, at least five times. Mass extinctions are defined by the loss of 75% of the biodiversity in a relatively short geological timeframe. These five mass extinctions, of which the most famous occurred at the end of Cretaceous and signed, the end of the Dinosaurs' era, were probably due to a synergy of cataclysmic events (Table 1).

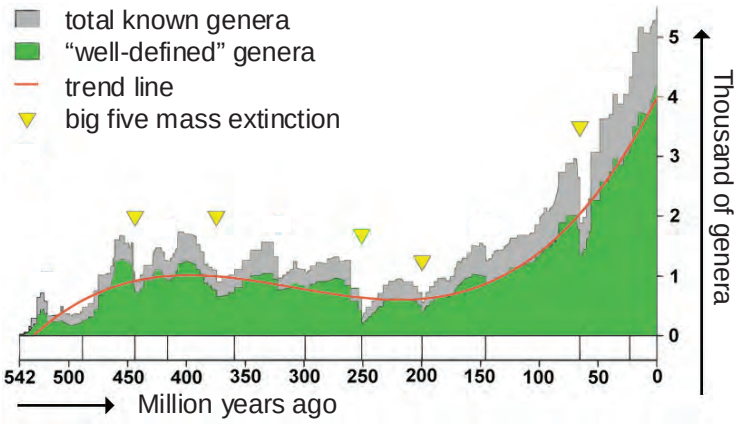


Figure 2. Representation of the global amount of biodiversity through time from the Phanerozoic to the Present based on fossil records (adapted from Rohde and Muller, 2005 and Wikipedia, 2013)

It is now widely accepted that current biodiversity is experiencing a massive wave of extinction possibly leading to a sixth mass extinction comparable to the other "Big Five" of past geological times, if nothing changes in human societies' behavior (Barnosky *et al.*, 2011). The causes of this decline in biodiversity are habitat loss and fragmentation, human monopolizing resources, introduction of non-native species and pathogens, and rapid climate warming (Wilcove *et al.*, 1998, MEA, 2005). For the first time, it is not extrinsic chemical, physical or astronomical phenomena that are responsible for this loss of biodiversity on Earth, but a single species, *Homo sapiens*. Our compelling responsibility in this biotic cataclysm, together with the fact that we are aware of the phenomenon should morally, urge to develop a reflection about the best measures to undertake in the perspective of reducing our impact on the decline of biodiversity.

Table 1. The 'Big Five' mass extinction events. From Barnosky *et al.* (2011)

Extinction event	**Causes of decline in biodiversity**
The Ordovician event ended ~443 Myr ago; 57% of genera were lost, an estimated 86% of species.	Onset of alternating glacial and interglacial episodes; repeated marine transgressions and regressions. Uplift and weathering of the Appalachians affecting atmospheric and ocean chemistry. Sequestration of CO_2.
The Devonian event ended ~359 Myr ago; 35% of genera were lost, an estimated 75% of species.	Global cooling (followed by global warming), possibly tied to the diversification of land plants, with associated weathering, paedogenesis, and the drawdown of global CO_2. Evidence for widespread deep-water anoxia and the spread of anoxic waters by transgressions. Timing and importance of bolide impacts still debated.
The Permian event ended ~251 Myr ago; 56% of genera were lost, an estimated 96% of species.	Siberian volcanism. Global warming. Spread of deep marine anoxic waters. Elevated H_2S and CO_2 concentrations in both marine and terrestrial realms. Ocean acidification. Evidence for a bolide impact still debated.
The Triassic event ended ~200 Myr ago; 47% of genera were lost, an estimated 80% of species.	Activity in the Central Atlantic Magmatic Province (CAMP) thought to have elevated atmospheric CO_2 levels, which increased global temperatures and led to a calcification crisis in the world oceans.
The Cretaceous event ended ~65 Myr ago; 40% of genera were lost, an estimated 76% of species.	A bolide impact in the Yucatán is thought to have led to a global cataclysm and caused rapid cooling. Preceding the impact, biota may have been declining owing to a variety of causes: Deccan volcanism contemporaneous with global warming; tectonic uplift altering biogeography and accelerating erosion, potentially contributing to ocean eutrophication and anoxic episodes. CO_2 spike just before extinction, drop during extinction.

WHY TO BE CONCERNED WITH ECOSYSTEMS PRESERVATION?

Seeking compromises between the apparent direct and short-term well-being of human societies and the preservation of ecosystem has not always been obvious. Why should we lower the value of our well-being for the conservation of biodiversity? This question implies that there is value in protecting threatened ecosystems. First, we can consider respecting the environment as an ethical value. In 1981, Paul Taylor introduced the concept of respecting our natural environment as a moral value similar in essence to respecting other people (Taylor, 1981). It sets the basis for life-centered environmental ethics where all living organisms, including human, have equal importance in the ecosystem functioning of the planet. Second, if the resources available in the ecosystems are not managed and preserved, they will inevitably be degraded in the long term and affect every component of global ecosystems. This was described as "the tragedy of the commons" (Hardin, 1968), i.e. when limited resources are available without any restrictions to anyone, the consequence is over-exploitation and resources exhaustion for everyone. Although this theory can be criticized and nuanced (e.g. Feeny *et al.*, 1990), it underlines that management efforts of all limited resources, biodiversity being part of it, should be, a target for a sustainable development of human societies. Finally, ecosystems provide services, i.e. add values to the capital they represent, e.g. by cleaning water, decomposing waste, pollinating crops, regulating climate or providing recreational benefits, their value being

thus more than just the capital they represent (Costanza *et al.*, 1997, MEA, 2005). Some attempts to quantify the values of the services provided by ecosystems (optimally preserved) have shown ratio possibly reaching 100:1 between pristine and over-exploited systems (Balmford *et al.*, 2002). Although these amounts may be subject to criticism or re-contextualization, evidences accumulate that there is an obvious intrinsic and direct benefit to preserve ecosystems on the long run.

This evidence leaded the international community to take measures for the conservation of biodiversity. In 1992, the Convention on Biological Diversity was opened for signatures at the Earth Summit of Rio De Janeiro and enforced in December 1993. The three main goals of the Convention are 1) conservation of biological diversity (or biodiversity), 2) sustainable use of its components and 3) fair and equitable sharing of benefits arising from genetic resources. Currently, all UN countries except the USA, Andorra and South Sudan have ratified the treaty. One of the most significant outputs of the Convention is the Biodiversity Assessment Reports published by the Millenium Ecosystem Assessment. Managing, preserving and planning biodiversity is now considered by the large majority as a central issue for a sustainable development of human society.

BIOLOGICAL INVASIONS: CAUSES AND CONSEQUENCES OF GLOBAL CHANGES

Biological invasions can be roughly defined as the process through which some species disperse outside of their historic distribution and cross natural geographic barriers with the help of mediated human activities (Richardson *et al.*, 2000, Fig. 3). These can be geographical, environmental or biotic barriers. Species are defined as invasive species when they present all the following characteristics: 1) they have been introduced, intentionally or not, as a result of human activity, 2) they are able to reproduce themselves and are thus sustaining exotic populations for several life cycles, and 3) they have the potential to spread over a considerable area (Richardson *et al.*, 2000). At the different phases of the invasion process, different terminologies are applied to qualify the species, reflecting the type of barriers they have crossed (Fig. 3). Additionally, invasive species are usually associated with nuisances to human society (Colautti and MacIsaac, 2004), either through direct or indirect effects. First, they can directly affect human health or well-being (Pyšek and Richardson, 2010). For example, Common Ragweed (*Ambrosia artemisiifolia*), a plant native from North America, has significantly impacted the rate of allergy in France since it has been introduced (Laaidi *et al.*, 2003). The spread of species like the Fire Ants (*Solenopsis invicta*), originating from South-America and now spreading rapidly in North America, China and Australia (Ascunce *et al.*, 2011), causes itchy stings to humans, making outdoor recreation unpleasant and hazardous. Second, invasive species can also affect human productivity. For example, the yellow star thistle (*Centaurea solstitalis*) now dominates millions of hectars in California grassland, an environment that used to be very productive (DeLong, 2002). The combination of direct impacts, monitoring, eradications and control of invasive species cost hundreds of billions dollars per year (Sinden *et al.*,

2004, Pimentel *et al.*, 2005, Lovell *et al.*, 2006). More generally, invasive species have significant impact on biodiversity and eventually on ecosystems. It is considered among the five main causes of the massive ongoing loss of biodiversity, combined with habitat destruction, pollution and overexploitation and climate change (Wilcove *et al.*, 1998, MEA, 2005). Invasive species may have disastrous impact, especially on isolated islands (e.g. Blackburn *et al.*, 2004, Medina *et al.*, 2011, Oppel *et al.*, 2011). For example, the Brown Tree Snake (*Boiga irregularis*), originating from Oceania, was boarded on a US military aircraft and introduced in Guam in the late 1940s. By the 1970s, released from its natural predators, it had spread on the entire island and had been responsible for the quasi complete extermination of all the island's native bird species (Lowe *et al.*, 2000). Invasive species may also have huge impact on the continents, where for example, the feral pig (*Sus scrofa*) can damage areas of native vegetation and spread weeds, disrupting ecological processes such as vegetation succession and local flora. Moreover, their diet can include juvenile land tortoises, sea turtles, sea birds and endemic reptiles (Lowe *et al.*, 2000). Invasive species not only impact species richness directly (Gaertner *et al.*, 2009), but also affect soils processes (Liao *et al.*, 2008), fire regimes (Brooks *et al.*, 2004) or trophic interactions (Traveset and Richardson, 2006) resulting in a chaotic cascade of possible causes able to deeply modify ecosystems.

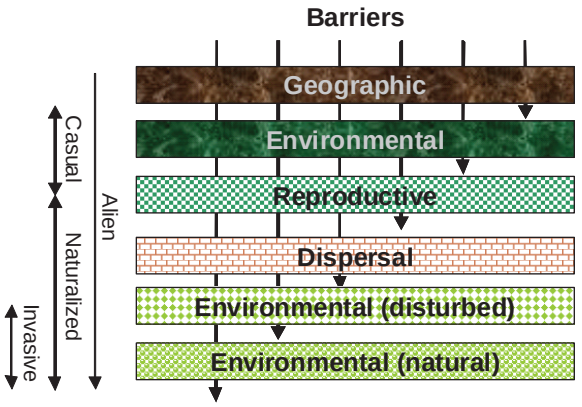


Figure 3. Different barriers that introduced species have to cross to be considered as invasive. The terminology used to qualify the species depends on stage of invasion (redrawn from Richardson *et al.*, 2000)

Biological invasions can also be favored by other global change components, such as habitat destruction, climate warming or nitrate deposition. For example, the large increase of evergreen broad-leaved species in southern Switzerland was shown to be highly correlated with the decreasing in the number of frost days per year (Walther *et al.*, 2002, Fig. 4). The complex interactions affecting the different components of global change even raise the question of whether invasive species are a cause or consequence of extinctions (Gurevitch and Padilla, 2004). Actually, all components originating from human are deeply intricate with mutual interactions and probably jointly induce effect on ecosystems (Fig. 5, Didham *et al.*, 2005, MacDougall and Turkington, 2005). Thus, it is necessary to study

biological invasions through the prism of global change in general, for a better contextualization and understanding of causes, processes and impacts. On the other hand, studying biological invasions may provide insightful information about the other components of global change.

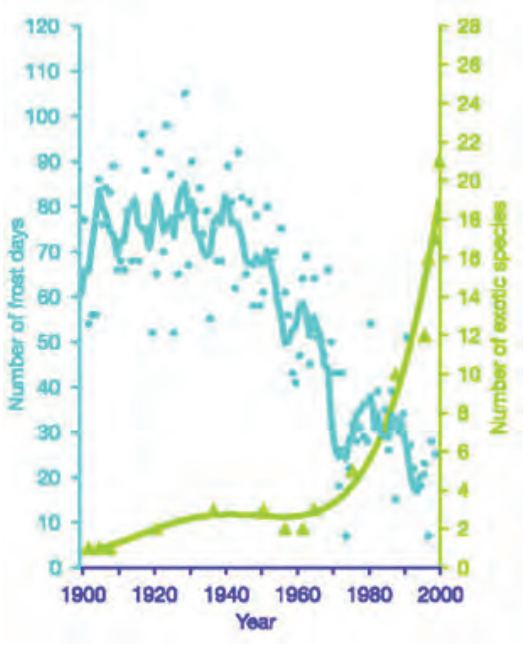


Figure 4. Correlation between the number of frost day per year and the number of non-native evergreen broadleaves species in southern Switzerland (from Walther *et al.*, 2002)

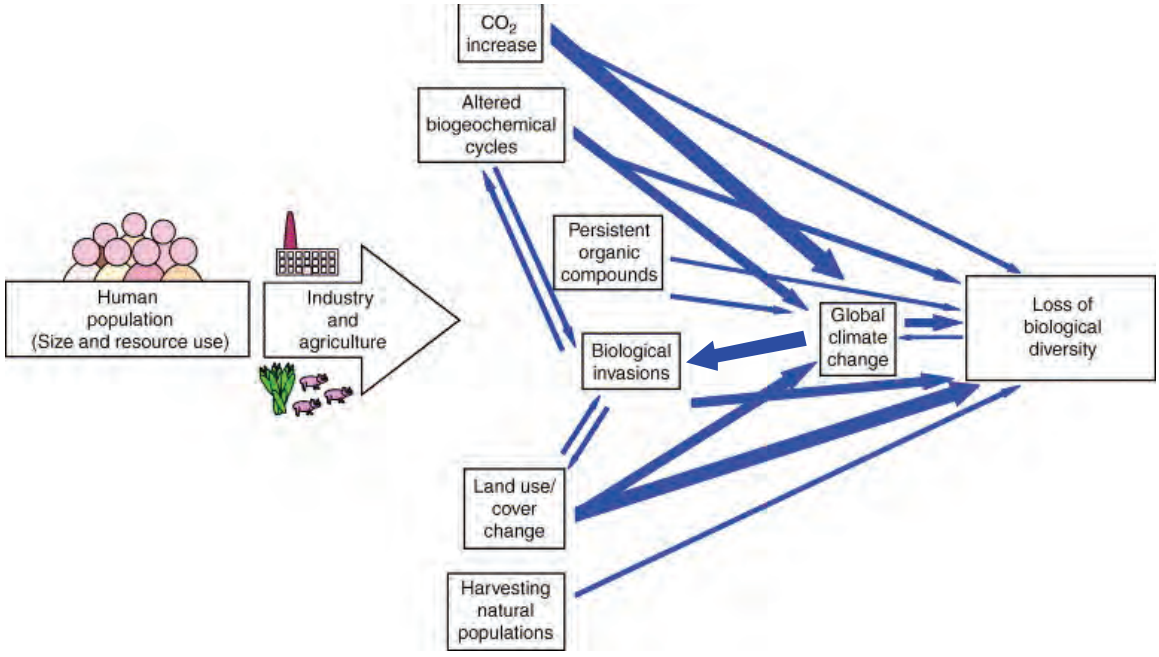


Figure 5. Biological invasions as only one of the numerous intricate components of global change (slightly adapted from Didham *et al.*, 2005)

INVASIVE SPECIES AS UNEXPECTED OPPORTUNITY TO UNDERSTAND SPECIES' RESPONSES TO GLOBAL CHANGE

Biological invasions are not only harmful, nuisances and part of the tragedy of the commons. They are also numerous unplanned and frequently, but imperfectly, replicated experiments using the Earth as a laboratory (Sax *et al.*, 2007). Among the numerous biogeographical concepts that have been highlighted by the study of biological invasions (e.g. dispersal limitation and equilibrium, sympatric speciation, Sax *et al.*, 2007), some may be particularly valuable to understand species' response to global change. For example, introduced species raised the awareness about the possibility that a single species might impact and transform a whole ecosystem (Vitousek *et al.*, 1987, D'Antonio and Vitousek, 1992). It also informed about how fast species can evolve and adapt to novel trophic interactions (Carroll *et al.*, 2005), or novel environments (Maron *et al.*, 2004, Broennimann *et al.*, 2007, Whitney and Gabler, 2008). Additionally, invaded species ranges can be used as independent data sets to test our ability to predict species distributions (Peterson and Vieglais, 2001, Thuiller *et al.*, 2005, Broennimann *et al.*, 2007) which is particularly important to validate predictions of future assemblages of biodiversity under climate warming (Beerling *et al.*, 1995). The necessity to understand biological invasions to better react and manage their impact, combined to the fascinating biological questions raised by their success lead biologist to develop a multitude of framework, approaches and studies to tackle this problem, one of the most popular being environmental niche modeling.

Note that the terminology about environmental niche modeling is debated (Kearney, 2006; Warren, 2012, McInerny and Etienne, 2012) and is more often qualified as ecological niche modeling. I choose to entitle my thesis with the term "environmental", because it describes a wider context than the "ecological" niche only. Typically in Chapter 3.1, I consider that when species niche is modeled using predictors such as urbanization density or the presence of river dams, this is beyond species "ecology" and rather deal with species "environment". However, the terminology Ecological Niche Modeling (ENM) is the most commonly found in the literature and is used in this thesis. An additional synonymous term, *species distribution modeling* (abbreviated SDM), judged as too vague by some authors, is also used in this thesis. Therefore, throughout this thesis, the abbreviation ENM and SDM will thus be considered as strictly synonymous, although ENM tend to be used in chapters that are more focused on the niche, whereas SDM is used in chapters dealing with the projection of the niche model into the geographical space.

ECOLOGICAL NICHE MODELLING TO STUDY INVASIVE SPECIES

Understanding species' geographic distributions requires understanding their environmental requirements. Joseph Grinell introduced the term "niche" to qualify the combination of environments in which a species can live (Grinnell, 1917), in other words the envelop of abiotic habitats characterizing where a species can be found. This concept

was later formalized by Hutchinson (1957) who defined the environmental niche as an n-dimensional hypervolume occupied by a species which characterized the main environmental variables influencing the species distribution. Additionally, Hutchinson raised an important distinction between the fundamental and the realized niche. The *fundamental* niche is delimited by the species' species' physiological tolerance along environmental gradients. It can be measured by investigating the species' ability to grow in a lab experiment where different levels of abiotic stress are applied to the organism (e.g. Oliver *et al.*, 2000). The *realized* niche corresponds to species' ability to grow in a given environment in the field, i.e. including all the biotic stressful factors such as predation, competition or diseases. It is generally agreed that the realized niche is a subset of the fundamental niche (Soberón, 2007, Fig. 6), unless positive interactions shape the distributions of species (Callaway *et al.*, 2002).

The realized niche is the pivotal concept underpinning ENM. ENM consists in relating species' observed occurrences to environmental factors through a statistical model in the form of response surfaces that can be used to predict the probability of occurrence of species, given particular conditions. If the environmental variables are spatially explicit, such as Geographical Information Systems' layers, it is then possible to assess the probability of occurrence (or suitability index) of every site of a given area by projecting the niche model in space (i.e. use it to predict species distribution, Fig. 7, Guisan and Zimmermann, 2000). Such models are thus a statistical formulation of Hutchinson's realized niche. Nevertheless, this latter may not be fully adequate to characterize ENM, in particular when species are still affected by ongoing dispersal. In this case, the BAM (for Biotic-Abiotic-Migration) diagram better reflects the niche depicted by ENM (Fig. 6, Soberón, 2007). According to this framework, a species distribution is limited by the distribution of available conditions in the study area, the environmental tolerances of the species (i.e. its fundamental niche), biotic interactions with other species and also by species ability to disperse. Since ENM is calibrated on the observed distribution of species, the modeled niche then implicitly includes all these limitations. Whether these limitations remain conserved through time and space is a debated topic and depends on many parameters, ranging from the species ability to retain their niche characteristics (niche conservatism), to the temporal and spatial scales at which the underpinning questions of the studies are asked (Wiens and Graham, 2005, Warren *et al.*, 2008, Peterson, 2011). All these limitations have to be considered with care.

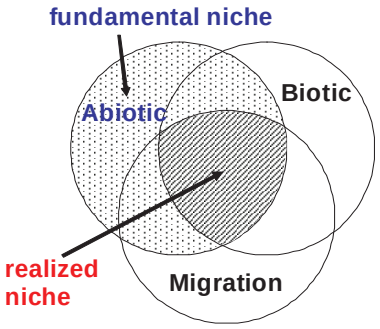


Figure 6. The BAM diagram. The realized niche modeled by ENM (hatched area) is the subset of the fundamental niche where biotic competition and dispersal migration are not limiting. The BAM diagram was proposed by (Soberón, 2007)

ENM has been widely used to test fundamental questions in ecological and evolutionary biology (Guisan and Thuiller, 2005, Richards *et al.*, 2007, Elith and Leathwick, 2009). Additionally, ENM is often one of the only possible way to optimize the management of a large number of species (i.e. biodiversity). They can be used in reserves design, for assessing the impact of climate, land use and other environmental changes on species distributions, for suggesting under-sampled sites suitable for rare species, for mapping potential species richness or for assessing species invasions and proliferation (Guisan and Thuiller, 2005, Guisan and Rahbek, 2011, Guisan *et al.*, 2013). In such cases, ENM commonly assumes that the realized niche modeled from the distribution of the species in the native range is conserved and can be transferred through space and time. The difficulty to test and validate models on independent data (Araujo *et al.*, 2005, Araujo and Guisan, 2006), especially when projected in the future, raised severe criticisms about the approach (Jeschke and Strayer, 2008), even leading some authors to postulate that ENM is not more informative than a null model (Beale *et al.*, 2008).

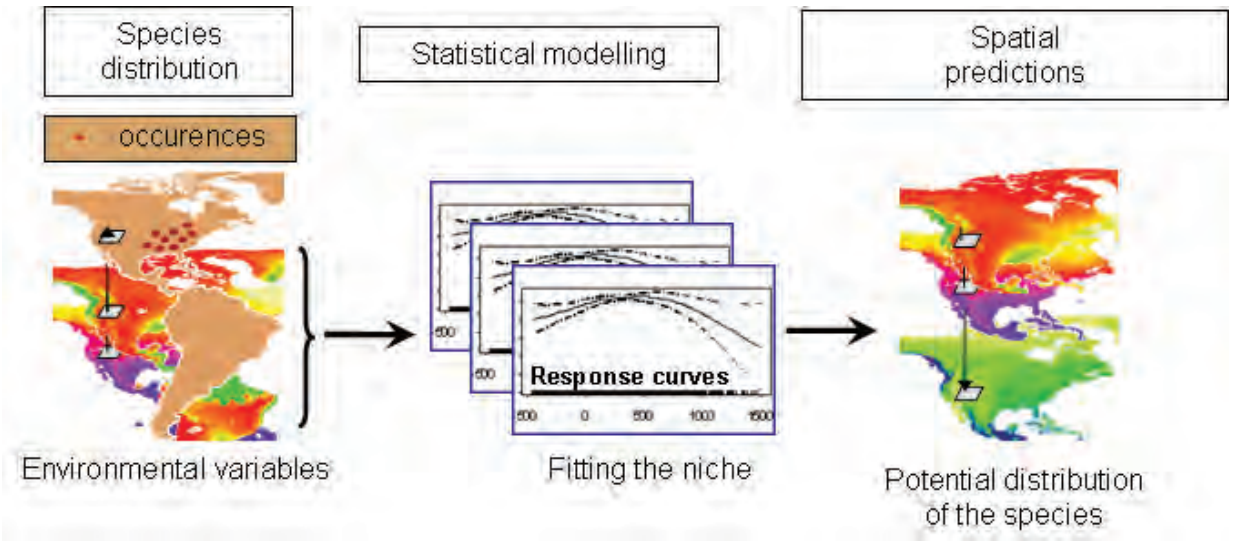


Figure 7. Scheme summarizing the environmental niche modeling approach. The main steps are the sampling of occurrences and the link to environmental variables, the fit of the niche in the environmental space and prediction of the model in space.

Invasive species are excellent model organisms to test how the species' niche can vary through space and time. They can inform about the dispersal rate and the time necessary for species to fill their niche, the species' ability to expand and colonize their realized niche once released from their native competitors and other enemies, their ability to expand their niche in climate that are non-existing (i.e. non-analog to) in the native range (e.g. pre-adaptation) or to inform about their evolutionary potential to adapt to novel environments. Understanding these factors is crucial to understand species' response to rapid global change. Studying biological invasions through ENM informs us not only about intrinsic characteristics of invasive species but also tells us about our ability to predict species distributions in the Anthropocene in general.

AIMS AND STRUCTURE OF THE THESIS

One of the major strength of ENM in our numerical era is the opportunity they offer to process the enormous amounts of species and climate data currently being collected at worldwide scale (Graham *et al.*, 2004). The general aim of this PhD thesis is to develop a unified framework that can be applied to many species and untangle the contradictory and often hardly comparable case studies study how introduced non-native species can be modeled with ENM to provide insightful information about invasion processes, to assess whether their niches are conserved between native and invaded ranges and ultimately to test our ability to predict biological invasions in a human-altered world. With this approach, I hope to shed light on more general patterns that will prove important to quantify, compare project niches, and ultimately draw generalities for the understanding of ecological and evolutionary processes. This thesis is built in three parts.

The first part tests for macro-climatic niche conservatism between native and invaded range of widely distributed plants and mammals. A first chapter **(1.1)** reviews the existing studies about niche shifts during biological invasions. It shows that many methodological disparities exist in the literature, potentially explaining the contradictory results among case studies. It proposes a new conceptual framework allowing a better interpretation of observed niche changes during biological invasions. A second chapter **(1.2)** uses this framework to seek for niche shifts due to competition release or rapid evolution in 50 Holarctic plant invaders. This analysis is extended in a third chapter **(1.3)** which tests the impact of bioclimatic variables selection for ENM transferability, with the aim of improving the predictions for Holarctic plant invaders under current and future climatic conditions. Finally, a fourth chapter **(1.4)** tests if the native climatic niche of 168 mammal's species predicts their success of introduction. As endothermic organisms, mammals are expected to be less impacted by climate but they are one of the rare organisms for which information about both successes and failures of introduction is available.

The second part of the manuscript aims to forecast biological invasions in a changing, human-altered world based on the new insights provided by the previous part. It starts with

a chapter **(2.1)** quantifying how climate warming may drive invasive plant species towards higher altitudes with currently preserved areas. The niche of 48 plant species, invasive in the Swiss and/or Australian Alps was modeled and their potential distribution was predicted under different climatic scenarios. It further analyses species distributional responses in regards to species traits and stressing factors which could differ between the two study areas. A second chapter **(2.2)** reveals the paradoxical and complex case of the Eurasian grapevine (*Vitis vinifera*) under Anthropocene changes. On the one hand, it has become one of the most widespread fruit crops in the world, but the wild ecotype is now endangered by habitat destruction and by the naturalized American *Vitis* species. The latter was introduced in the 19th century as rootstocks in the vineyards to protect viticulture against the devastative spread of *Phylloxera*. ENM was used to quantify the niche overlap between these three conflicting taxonomic unit in order to assess current and future threats on the already endangered wild grapevine.

The third and final part studies the niche and distribution of invasive species beyond the macro-climate scale. As reflected by the previous chapters of my thesis dealing with coarse resolutions from regional to global scales, few studies have been conducted at finer scale, where conservation planning ultimately takes place and where non-climatic factors become more important. The final chapter **(3.1)** presents the changes at the population scale of five invasive plant species distributed along an 18 km section of a Swiss river including some important protected alluvial areas, over an 11 years period. Three species showed an overall dramatic increase of the number of populations and individuals along the river section. ENM was built at a very high resolution (1 m), revealing the preponderant role of light-availability, topography, river and human disturbances. The study highlights the importance of anticipating invasions at extremely fine scale to allow better conservation prioritizations.

At the end of the document, I attached several appendices where I was involved as a co-author and that aim to widen the use of the tools and the problematic developed during this thesis (see the *Synthesis and Moving Forward* section).

REFERENCES

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

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

Ascunce, M.S., Yang, C.-C., Oakey, J., Calcaterra, L., Wu, W.-J., Shih, C.-J., Goudet, J., Ross, K.G. & Shoemaker, D. (2011) Global Invasion History of the Fire Ant *Solenopsis invicta*. *Science*, **331,** 1066-1068.

Augustin, L., Barbante, C., Barnes, P.R., Barnola, J.M., Bigler, M., Castellano, E., Cattani, O., Chappellaz, J., Dahl-Jensen, D. & Delmonte, B. (2004) Eight glacial cycles from an Antarctic ice core. *Nature*, **429,** 623-628.

Balmford, A., Bruner, A., Cooper, P., Costanza, R., Farber, S., Green, R.E., Jenkins, M., Jefferiss, P., Jessamy, V., Madden, J., Munro, K., Myers, N., Naeem, S., Paavola, J., Rayment, M., Rosendo, S.,

Roughgarden, J., Trumper, K. & Turner, R.K. (2002) Economic Reasons for Conserving Wild Nature. *Science*, **297,** 950-953.

Barnosky, A.D., Matzke, N., Tomiya, S., Wogan, G.O.U., Swartz, B., Quental, T.B., Marshall, C., McGuire, J.L., Lindsey, E.L., Maguire, K.C., Mersey, B. & Ferrer, E.A. (2011) Has the Earth's sixth mass extinction already arrived? *Nature*, **471,** 51-57.

Beale, C.M., Lennon, J.J. & Gimona, A. (2008) Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences*, **105,** 14908-14912.

Beerling, D.J., Huntley, B. & Bailey, J.P. (1995) Climate and the distribution of *Fallopia japonica*: use of an introduced species to test the predictive capacity of response surfaces. Journal of Vegetation Science, **6,** 269-282.

Blackburn, T.M., Cassey, P., Duncan, R.P., Evans, K.L. & Gaston, K.J. (2004) Avian extinction and mammalian introductions on oceanic islands. *Science*, **305,** 1955-1958.

Broecker, W.S. & Stocker, T.F. (2006) The Holocene CO2 rise: Anthropogenic or natural? *Eos, Transactions American Geophysical Union*, **87,** 27-27.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10,** 701-709.

Brooks, M.L., D'Antonio, C.M., Richardson, D.M., Grace, J.B., Keeley, J.E., DiTomaso, J.M., Hobbs, R.J., Pellant, M. & Pyke, D. (2004) Effects of invasive alien plants on fire regimes. *BioScience*, **54,** 677-688.

Callaway, R.M., Brooker, R., Choler, P., Kikvidze, Z., Lortie, C.J., Michalet, R., Paolini, L., Pugnaire, F.I., Newingham, B. & Aschehoug, E.T. (2002) Positive interactions among alpine plants increase with stress. *Nature*, **417,** 844-848.

Carroll, S.P., Loye, J.E., Dingle, H., Mathieson, M., Famula, T.R. & Zalucki, M.P. (2005) And the beak shall inherits evolution in response to invasion. *Ecology Letters*, **8,** 944-951.

Colautti, R.I. & MacIsaac, H.J. (2004) A neutral terminology to define 'invasive' species. *Diversity and Distributions*, **10,** 135-141.

Costanza, R., dArge, R., deGroot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., Oneill, R.V., Paruelo, J., Raskin, R.G., Sutton, P. & vandenBelt, M. (1997) The value of the world's ecosystem services and natural capital. *Nature*, **387,** 253-260.

Crutzen, P.J. (2002) Geology of mankind. *Nature*, **415,** 23-23.

D'Antonio, C.M. & Vitousek, P.M. (1992) Biological invasions by exotic grasses, the grass/fire cycle, and global change. *Annual review of ecology and systematics*, **23,** 63-87.

DeLong, B. (2002) http://www.j-bradford-delong.net/movable_type/2003_archives/000638.html.

Didham, R.K., Tylianakis, J.M., Hutchison, M.A., Ewers, R.M. & Gemmell, N.J. (2005) Are invasive species the drivers of ecological change*? Trends in Ecology & Evolution*, **20,** 470-474.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, **40,** 677-697.

Feeny, D., Berkes, F., McCay, B.J. & Acheson, J.M. (1990) The tragedy of the commons: twenty-two years later. *Human ecology*, **18,** 1-19.

Gaertner, M., Den Breeyen, A., Hui, C. & Richardson, D.M. (2009) Impacts of alien plant invasions on species richness in Mediterranean-type ecosystems: a meta-analysis. *Progress in Physical Geography*, **33,** 319-338.

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution*, **19,** 497-503.

Grinnell, J. (1917) The niche-relationships of the California Thrasher. The Auk, **34,** 427-433.

Guisan, A. & Rahbek, C. (2011) SESAM - a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography*, **38,** 1433-1444.

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

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16,** 1424-1435.

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

Gurevitch, J. & Padilla, D.K. (2004) Are invasive species a major cause of extinctions? *Trends in Ecology & Evolution*, **19,** 470-474.

Hardin, G. (1968) The Tragedy of the Commons. *Science*, **162,** 1243-1248.

Holdaway, R.N. & Jacomb, C. (2000) Rapid Extinction of the Moas (*Aves*: *Dinornithiformes*): Model, Test, and Implications. *Science*, **287,** 2250-2254.

Hutchinson, G.E. (1957) Population Studies - Animal Ecology and Demography - Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22,** 415-427.

Jeschke, J.M. & Strayer, D.L. (2008) Usefulness of bioclimatic models for studying climate change and invasive species. *Annals of the New York Academy of Sciences*, **1134,** 1-24.

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

Kump, L.R., Kasting, J.F. & Crane, R.G. (2010*) The earth system*. Prentice Hall, San Francisco, CA.

Laaidi, M., Laaidi, K., Besancenot, J.-P. & Thibaudon, M. (2003) Ragweed in France: an invasive plant and its allergenic pollen. *Annals of Allergy, Asthma & Immunology*, **91,** 195-201.

Liao, C., Peng, R., Luo, Y., Zhou, X., Wu, X., Fang, C., Chen, J. & Li, B. (2008) Altered ecosystem carbon and nitrogen cycles by plant invasion: a meta-analysis. *New Phytologist*, **177,** 706-714.

Lovell, S.J., Stone, S.F. & Fernandez, L. (2006) The economic impacts of aquatic invasive species: a review of the literature. *Agricultural and Resource Economics Review*, **35,** 195.

Lowe, S., Browne, M., Boudjelas, S. & De Poorter, M. (2000*) 100 of the world's worst invasive alien species: a selection from the global invasive species database*. Invasive Species Specialist Group Auckland, New Zealand.

MacDougall, A.S. & Turkington, R. (2005) Are invasive species the drivers or passengers of change in degraded ecosystems? *Ecology*, **86,** 42-55.

Maron, J.L., Vila, M., Bommarco, R., Elmendorf, S. & Beardsley, P. (2004) Rapid evolution of an invasive plant. *Ecological Monographs*, **74,** 261-280.

Martin, P.S. (1973) The Discovery of America: The first Americans may have swept the Western Hemisphere and decimated its fauna within 1000 years. *Science*, **179,** 969-974.

McInerny, G.J. & Etienne, R.S. (2012) 'Niche' or 'distribution' modelling? A response to Warren. *Trends in Ecology & Evolution*.

McNeill, J.R. (2000) *Something New Under the Sun: An Environmental History of the Twentieth-Century World (The Global Century Series)*. WW Norton & Company, New York, London.

MEA (2005) *Millennium Ecosystems Assessment: Ecosystems and Human Well-Being.* Island press, Washington, DC.

Medina, F.M., Bonnaud, E., Vidal, E., Tershy, B.R., Zavaleta, E.S., Josh Donlan, C., Keitt, B.S., Corre, M., Horwath, S.V. & Nogales, M. (2011) A global review of the impacts of invasive cats on island endangered vertebrates. *Global Change Biology*, **17,** 3503-3510.

Oliver, M., Velten, J. & Wood, A. (2000) Bryophytes as experimental models for the study of environmental stress tolerance: Tortula ruralis and desiccation-tolerance in mosses. *Plant Ecology*, **151,** 73-84.

Oppel, S., Beaven, B.M., Bolton, M., Vickery, J. & Bodey, T.W. (2011) Eradication of invasive mammals on islands inhabited by humans and domestic animals. *Conservation Biology*, **25,** 232-240.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, **38,** 817-827.

Peterson, A.T. & Vieglais, D.A. (2001) Predicting Species Invasions Using Ecological Niche Modeling: New Approaches from Bioinformatics Attack a Pressing Problem: A new approach to ecological niche modeling, based on new tools drawn from biodiversity informatics, is applied to the challenge of predicting potential species' invasions. *BioScience*, **51,** 363-371.

Pimentel, D., Zuniga, R. & Morrison, D. (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecological Economics*, **52,** 273-288.

Pyne, S.J. (1995) *World fire: the culture of fire on earth*. University of Washington Press, Seattle.

Pyšek, P. & Richardson, D.M. (2010) Invasive Species, Environmental Change and Management, and Health. *Annual Review of Environment and Resources*, **35,** 25-55.

Richards, C.L., Carstens, B.C. & Knowles, L.L. (2007) Distribution modelling and statistical phylogeography: an integrative framework for generating and testing alternative biogeographical hypotheses. *Journal of Biogeography*, **34,** 1833-1845.

Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions*, **6,** 93-107.

Roberts, R.G., Flannery, T.F., Ayliffe, L.K., Yoshida, H., Olley, J.M., Prideaux, G.J., Laslett, G.M., Baynes, A., Smith, M.A., Jones, R. & Smith, B.L. (2001) new ages for the last australian megafauna: continent-wide extinction about 46,000 years ago. *Science*, **292,** 1888-1892.

Rohde, R.A. & Muller, R.A. (2005) Cycles in fossil diversity. *Nature*, **434,** 208-210.

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S.D., Grosberg, R.K., HastingS, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I. & Rice, W.R. (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution*, **22,** 465-471.

Sinden, J., Jones, R., Hester, S., Odom, D., Kalisch, C., James, R., Cacho, O. & Griffith, G. (2004) The economic impact of weeds in Australia. *Technical Series*, **8**.

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

Steffen, W., Crutzen, P.J. & McNeill, J.R. (2007) The Anthropocene: are humans now overwhelming the great forces of nature. *AMBIO: A Journal of the Human Environment*, **36,** 614-621.

Taylor, P.W. (1981) The Ethics of Respect for Nature. *Environmental Ethics*, **3,** 197-218.

Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (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.

Traveset, A. & Richardson, D.M. (2006) Biological invasions as disruptors of plant reproductive mutualisms. *Trends in Ecology & Evolution*, **21,** 208-216.

Vitousek, P.M., Walker, L.R., Whiteaker, L.D., Mueller-Dombois, D. & Matson, P.A. (1987) Biological invasion by Myrica faya alters ecosystem development in Hawaii. *Science*, **238,** 802-804.

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

Warren, D.L. (2012) In defense of "niche modeling". *Trends in Ecology & Evolution*, **27,** 497-500.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62,** 2868-2883.

Whitney, K.D. & Gabler, C.A. (2008) Rapid evolution in introduced species' invasive traits and recipient communities: challenges for predicting invasive potential. *Diversity and Distributions*, **14,** 569-580.

Wiens, J.J. & Graham, C.H. (2005) Niche conservatism: integrating evolution, ecology, and conservation biology. *Annual review of ecology, evolution, and systematics***,** 519-539.

Wilcove, D.S., Rothstein, D., Dubow, J., Phillips, A. & Losos, E. (1998) Quantifying threats to imperiled species in the United States. *BioScience*, **48,** 607-615.

PART I

Test for niche conservatism between native and invaded ranges

CHAPTER 1.1

Unifying niche shift studies: insights from biological invasions

Antoine Guisan[1], Blaise Petitpierre[1], Olivier Broennimann[1], Curtis Daehler[2] & Christoph Kueffer[3]

[1] Department of Ecology & Evolution, University of Lausanne, 1015 Lausanne, Switzerland

[2] Department of Botany, University of Hawaii at Manoa, 3190 Maile Way, Honolulu, HI 96822, USA

[3] Institute of Integrative Biology, ETH Zürich, Universitätstrasse 16, 8092 Zürich, Switzerland

This paper accepted in *Trends in Ecology and Evolution*

My contribution to the project: this paper was initially started as a response to a technical comment about chapter 1.2. I contributed in the framework presented in this paper, participated to the literature review, conceptualized and drew the Figure 2 and S1 and took part to the writing process.

ABSTRACT

Assessing whether the climatic niche of a species may change between different geographic areas or time periods has become increasingly important in the context of ongoing global change. However, approaches and findings have remained largely controversial so far, calling for a unification of methods. Here, we build on a review of empirical studies of invasion to formalize a unifying framework that decomposes niche change into unfilling, stability and expansion situations, taking both a pooled-range and range-specific perspective on the niche, while accounting for climatic availability and climatic analogy. This framework provides new insights into the nature of climate niche shifts and our ability to anticipate invasions, and may help guiding the design of experiments for assessing causes of niche changes.

NICHE SHIFTS DURING INVASIONS: SETTING THE SCENE

How climate determines the distribution of species is a classic question in ecology, closely tied to Hutchinson's (1957) concept of the environmental niche, and still a major research topic (Soberón & Nakamura, 2009, Peterson *et al.*, 2011). Although, for some species, it has become possible to determine the fundamental climatic niche based on physiological information and a mechanistic understanding (e.g. Kearney & Porter, 2009), for most species only the realized climate niche can realistically be estimated through empirical studies (Guisan & Thuiller, 2005). With global change, it has become increasingly important not only to describe species' climate niches but also to understand whether these can change rapidly (niche shifts) or not (niche conservatism) between different geographic areas or time periods (Peterson, 2011; Fig. 1). The primary approach to investigating climatic niches in space and time has been to analyse climatic conditions across a species' distributional ranges and/or over time (Pearman *et al.*, 2008a).

As already understood by Charles Elton (1958), biological invasions offer a unique opportunity to study how species colonize new environments (Sax *et al.*, 2007; Richardson & Pysek, 2008, Kueffer *et al.*, 2013), and whether they retain their climatic niche in a new range (Pearman *et al.*, 2008a). Addressing this question has proved important in recent years as a test of our capacity to use climate matching to assess invasion risks by exotic species at transnational scales (Venette *et al.*, 2010, Guisan *et al.*, 2013), in particular when using ecological niche models (ENMs), which rely heavily on climatic niche conservation between native and exotic ranges (Pearman *et al.*, 2008a, Colwell & Rangel, 2009, Peterson, 2011). Do a majority of species retain their native climatic niche when introduced elsewhere? The answer to this question is fundamental because it informs both theoretical and applied ecology, but approaches have diverged and findings have remained largely controversial so far (Table S1, supplementary material) (Pearman *et al.*, 2008*a*; Peterson, 2011).

Evidence exists both for and against climatic niche conservatism during invasions. A recent large-scale survey of 50 Holarctic terrestrial plant invaders concluded that climatic

niche shifts are rare overall between the native and invaded ranges, and therefore models can usefully predict invasion in the exotic range (Petitpierre *et al.*, 2012).) The same conclusions were reached for birds (Strubbe *et al.*, 2013) and other groups (see Peterson (2011), Table S1, suppl. mat.). But the assumption of niche conservatism was also challenged by evidence of climatic niches shifting during invasions (e.g. Broennimann *et al.*, 2007, Fitzpatrick *et al.*, 2007, Rödder & Lötters, 2009, Medley, 2010, Lauzeral *et al.*, 2011, Table S1), potentially hampering predictions in the new range. Contrasting evidence of niche dynamics during invasions, and particularly of the frequency of niche shifts (i.e. of centroid and/or limits; see Fig. 1) among various taxonomic groups, thus coexist in the literature (about 50% shifts / 42% no-shifts and 8% no-conclusion in Table S1). This contrasting evidence may, however, correspond to different types of niche changes, biological and/or methodological study contexts, data types, species characteristics or methods being used (Peterson & Nakazawa, 2008, Mandle *et al.*, 2010, Peterson, 2011, Soberón & Peterson, 2011, Broennimann *et al.*, 2012, Table S1), which confounding effects prevent sound interpretation of the possible mechanisms behind niche changes. Unification of the analytic context and metrics used, and a well-balanced comparison across different species, taxonomic groups, environmental spaces and geographic areas (Pearman *et al.*, 2008*a*, Kueffer *et al.*, 2013), may contribute to reconcile conflicting evidence in observational studies of biological invasions requires.

Here, we build on a review of niche changes reported in empirical invasion studies (Table S1) to formalize a new framework that unifies the analytical context (Box 1, Figs 1 and 2), clarifies the role of the niche-biotope duality (Box 2, Colwell & Rangel, 2009; Soberón & Nakamura, 2009), and helps to identify potential factors influencing niche change between ranges. The central idea of this framework is to decompose a niche comparison between native and exotic ranges into its three basic components: niche unfilling, niche stability and niche expansion (Box 3, Fig. 2, Petitpierre *et al.*, 2012). We present these elements and discuss them along with the importance of taking into account the available environment, distinguishing analog from non-analog climatic conditions between ranges (Box 4), and accounting for niche factors and niche dynamics at finer resolution. We conclude with recommendations on using the proposed framework for future niche change studies.

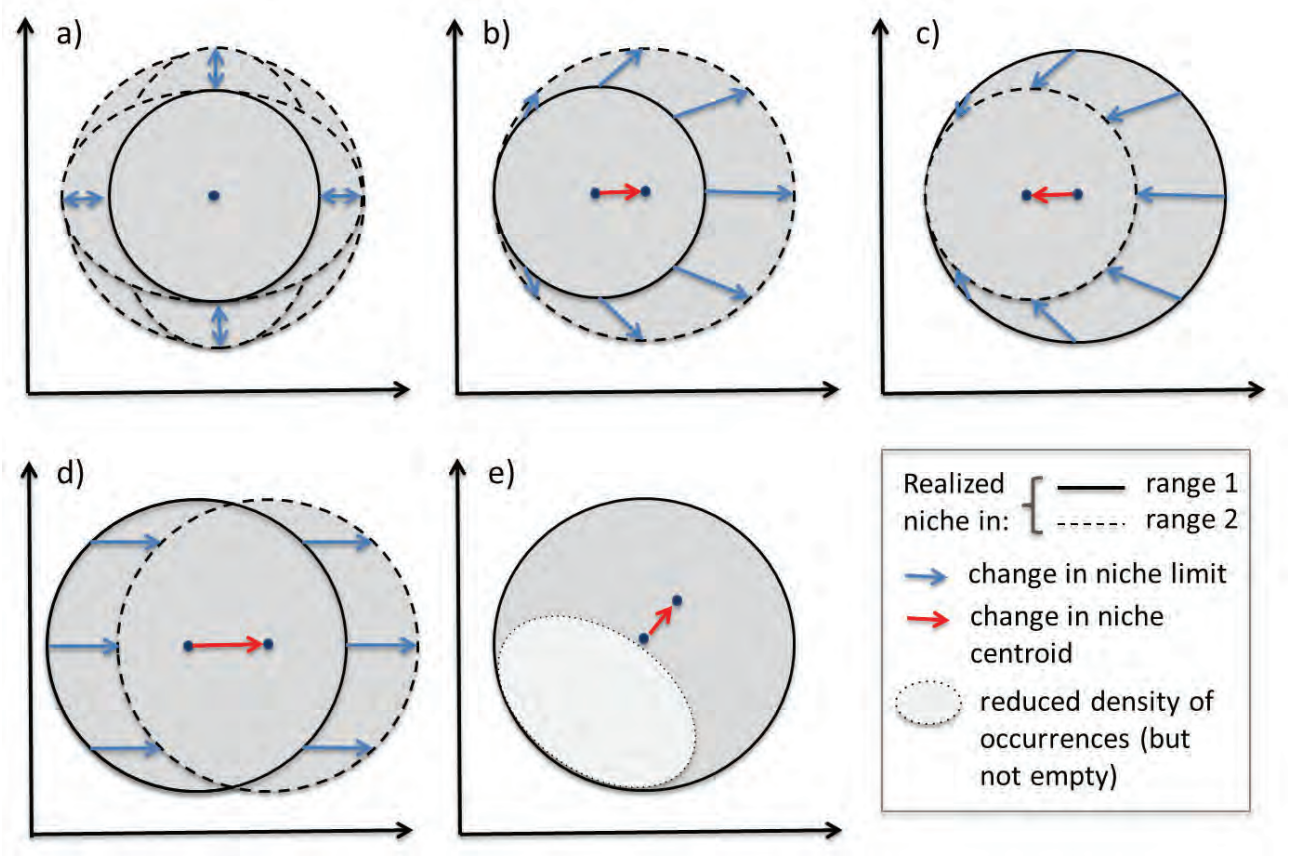


Figure 1: Theoretical scenarios of realized niche changes in space (e.g. following invasions) or time (e.g. under climate change). Change of : (i) the niche envelop (expansion or contraction) without change of the niche centroid, due to symetric niche change, i.e. in two opposite (a) or all directions in climatic space; (ii) the niche centroid with expansion (b, c) or displacement (d) of part of or the whole niche envelop , or (iii) the niche centroid only, due to a change of the density of occurrences within the same niche envelop in climatic space (e). The latter case would result in stability (no change) in figure 2. Observed changes are likely to be combinations of these cases.

Box 1: The analytical context for quantifying niche shifts

Assessing niche change between ranges is generally done by considering a species native in one area (its native range) and invading another (or several other) biogeographically separated area (the exotic range; e.g. Petitpierre *et al.*, 2012). This context could similarly apply to the same species in two (or more) time periods (e.g. Maiorano *et al.*, 2012). Regions large enough to include the entire (or large parts of) the native and exotic species' geographic distributions are usually considered for comparison. The choice of these areas will strongly condition the niche-biotope duality (Box 2), and thus the available environments (Fig. 2, Box 3), and ultimately the quantification of niche changes (Soberón & Nakamura, 2009, Soberón & Peterson, 2011). Optimally, the studied ranges should encompass the species' complete geographic distribution in the native and introduced ranges that could potentially be reached by a species given its dispersal ability, i.e. the accessible areas (Barve *et al.*, 2011). In practice we recommend defining areas with ecological relevance, such as biomes or ecoregions, and using species data (atlas or occurrences) well representing the focal species' range. The full multi-dimensional set of environmental conditions observed in one area/time period is the realized environment (Box 2, Box 3, Jackson & Overpeck, 2000, Ackerly, 2003) and the envelop of conditions where the species is observed represents its realized environmental niche (Box 2, Araujo & Guisan, 2006, Soberón & Nakamura, 2009).

NICHE CHANGES AND ASSOCIATED METRICS

Which niche is measured from field observations?

The realized climatic niche quantified from field observations is determined by biotic constraints on the fundamental eco-physiological niche, population dynamics (e.g. source-sink dynamics) and dispersal limitations (i.e. accessibility; Box 2, Pulliam, 2000, Soberón, 2007, Barve *et al.*, 2011), but it is also constrained by the availability of the environment in the areas (Box 4) at the timescale considered in the study (i.e. some conditions can be available at one time in one area, but not earlier or later, Jackson & Overpeck, 2000, Mandle *et al.*, 2010). A change in this realized niche can thus result from adaptive evolution occurring in the colonized range (Sax *et al.*, 2007, Alexander & Edwards, 2010) or from changes in biotic interactions, dispersal limitations, or from pre-adaptation to conditions not (anymore) available in the initial range at the time of the study but available in the colonized range (Pearman *et al.*, 2008*a*). Hereafter, we consider a niche shift as any change of the realized niche, i.e. the niche as measured by climatic characteristics at sites of species occurrence in the field. It thus includes implicitly any potential change of the fundamental niche, although with such empirical data, a change caused by evolution of physiological tolerance cannot be differentiated from a change due to other factors (Broennimann *et al.*, 2007, Soberón & Peterson, 2011).

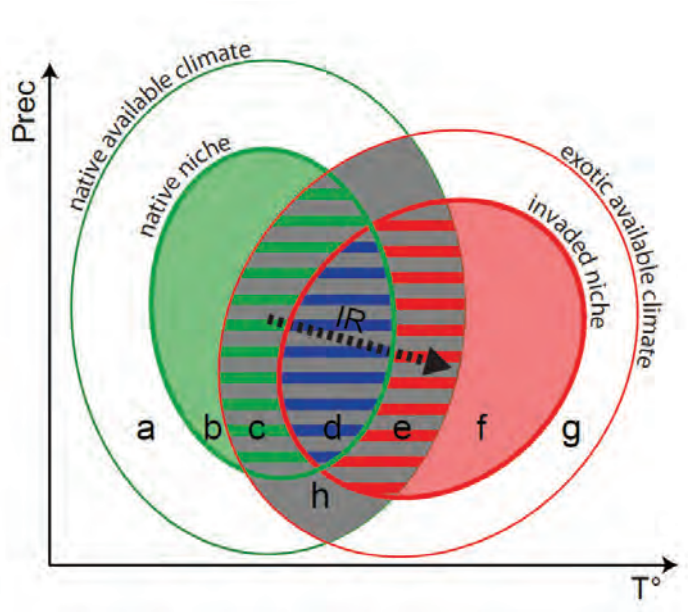
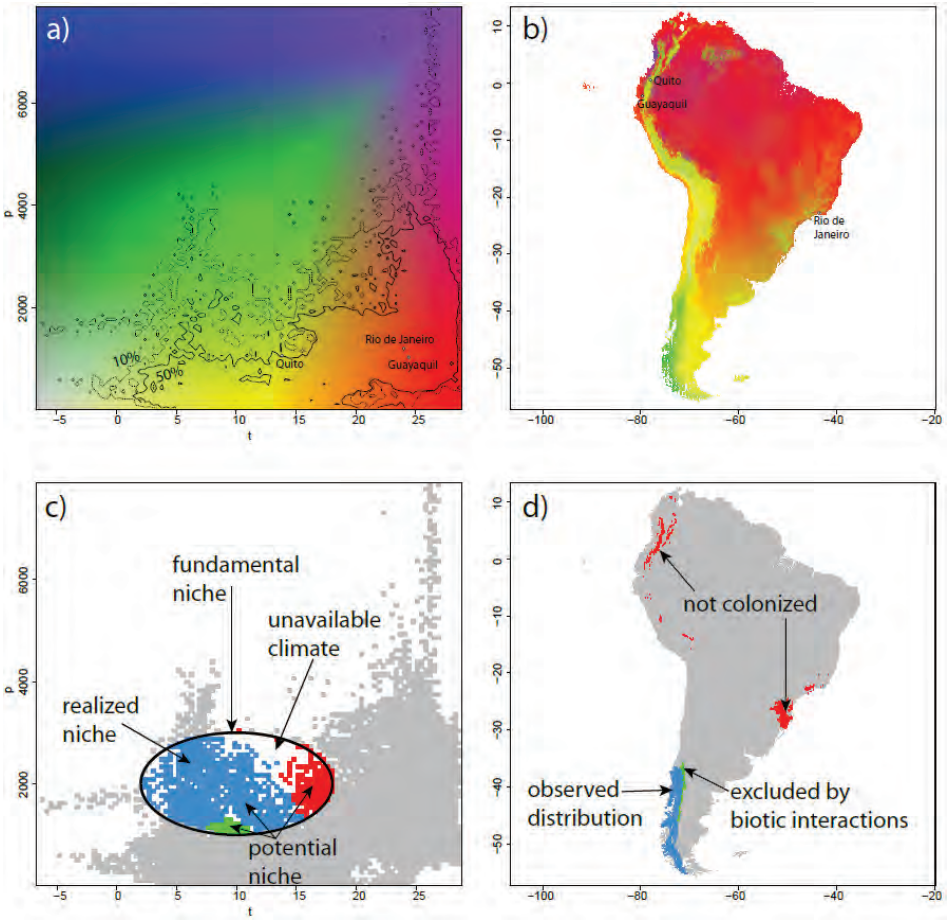


Figure 2: Schematic 2-dimensional representation of the indices of niche change (unfilling, stability and expansion) presented in Petitpierre *et al.* (2012) (see definitions in Box 3). Solid thin lines show the density of available environments (Box 4) in the native range (in green) and in the invaded range (in red). The gray area shows the most frequent environments common to both ranges (i.e. analog environments). The green and red thick lines show respectively the native and the invaded niches. Niche unfilling (U), stability (Se) and expansion (E) are shown respectively with green, blue and red hatched surfaces inside analog environments. The definition of a niche shift using the change of niche centroid only (inertia ratio, IR) is shown with a thick dotted arrow. In this context, the lower-case letters represent similar features in both graphs: a. available conditions in the native range, outside of the native niche and non-analog to the invaded range.. b. Conditions inside of the native niche but non-analog to the invaded range. c. Unfilling, i.e. conditions inside of the native niche but outside the invaded niche , possibly due to recent introduction combined with ongoing dispersal of the exotic species, which should at term fill these conditions. d. Niche stability, i.e. conditions filled in both native and invaded range. e. Niche expansion, i.e. conditions inside the invaded niche but outside the native one, due to ecological or evolutionary change in the invaded range. f. Conditions inside of the invasive niche but non-analog to the native range. g. Available conditions in the invaded range but outside of the invasive niche and non-analog to the native range. h. Analog conditions between the native and invaded ranges.

Box 2: Hutchinson's niche-biotope duality

It is important to recall the niche (environmental space) versus biotope (geographic space) duality framework described by G.E. Hutchinson (see Colwell & Rangel, 2009, Glossary). This duality means that there is no direct match between the topology of the niche space and the geographic distribution of a species (see figure below). The same combination of climate factors (colors in panel a of the figure) can occur in one or several localities in geographic space (same colors in panel b), and locations close in environmental space can be far apart geographically and vice versa (Soberón & Nakamura, 2009). For instance, in South America, the cities of Quito and Guayaquil are close to each other but climatically far away, whereas Guayaquil and Rio are geographically far but climatically close (see figure below). Interpretation of niches and distributions of species thus requires careful screening of both spaces jointly (see figure a,c vs. b,d), with special attention to issues of dispersal limitations, biotic interactions and available environmental conditions (Colwell & Rangel, 2009). Blank areas in panel c of the figure below represent environments that are not available within the geographic range considered (here South America). The intersection of the available environment and the fundamental (i.e. physiological) limits of a species define its potential niche (Jackson & Overpeck, 2000, Soberón & Nakamura, 2009). Parts of this potential niche can be unoccupied by the species because of dispersal limitations (i.e. red areas in panels c and d) or exclusion by biotic interactions (i.e. green parts in panels c and d). As the potential niche rests on the hypothetical quantification of the fundamental niche (see Sax *et al.*, 2013,)), whereas we focus mainly on the realized niche here, we do not expand further on this concept. For a full theoretical development of the concepts and definitions of niches and distributional areas, with formal abbreviations, see Soberón & Nakamura (2009) and Peterson *et al.* (2011).



Hutchinson's duality framework, inspired by Soberón and Nakamura (2009) and Soberón & Peterson (2011). The fundamental niche ellipse pictured in panel c) is theoretical (artificially created) and could not be derived from field observations.

Two main approaches to quantify niche changes

Two main approaches have been used so far to compare niches between ranges, based on direct observations or on model predictions (Broennimann *et al.*, 2012, Fig. 3, Table S1). The first approach uses observations directly and compares the difference in environmental attributes of the sites where the species occurs between the native and exotic ranges in environmental space. This comparison can be done either through univariate (e.g. Lauzeral *et al.*, 2011) or multivariate tests (e.g. in a reduced PCA space, Broennimann *et al.*, 2007, Fig. 3a). Such a direct approach does not rely on any underlying model that relates the occurrences to the environment. The approach can be considerably improved by calculating smooth densities of species occurrences in a gridded environmental space, as a way to avoid unrealistic 'holes' in a niche due to low sampling effort (Broennimann *et al.*, 2012) (see also Guisan *et al.*, 2012, Webber *et al.*, 2012). The second approach relies instead on the outcomes of ecological niche models (ENMs, Peterson *et al.*, 2011); also called species distribution models, (SDMs, Guisan & Thuiller, 2005), and compares the overlap of reciprocal predictions of geographic distributions (i.e. predicting the invaded distribution with the model fitted in the native range, and vice-versa), usually comparing in the exotic range the two predictions by the models fitted in each range (Fitzpatrick *et al.*, 2007, Warren *et al.*, 2008, Warren *et al.*, 2010, Fig. 3b). Specialized software has been developed for niche comparisons based on the ENM approach (ENMTool, Warren *et al.*, 2010). Comparative analyses with virtual species, for which distributions and niche overlap are known, showed that the first approach (ordination) quantified niche overlap overall more accurately than the second (ENM, Broennimann *et al.*, 2012); however, the ordination approach provides a mathematically less formalized representation of the niche and is less able to optimize the weighting of the different environmental factors based on their relevance for a species' ecology. The ENM approach is particularly useful to assess ENM transferability between ranges (Randin *et al.*, 2006). Thus, although both approaches have strengths and weaknesses (Broennimann *et al.*, 2012), comparisons of niche change results between studies (meta-analyses) should include preferentially those based on ordinations, and at least make clear which approach was used (see Table S1).

Different components of niche change: centroid shift, overlap, expansion and unfilling

From either of these approaches, different niche change metrics can be calculated, at two levels of analyses – pooled ranges and range-specific (Box 3). The most commonly used metrics so far measure either a shift of the niche centroid, C (mean position; e.g. using Euclidean distance, Broennimann *et al.*, 2007), or a change in the overlap, O, between the two niches (e.g. using Schoener's D, Warren *et al.*, 2008) or minimum convex polygons (Gallagher *et al.*, 2010), and they are usually calculated in relation to the entire realized niche between two ranges (i.e. pooled; Box 3). However, a niche change detected in one of these two ways can result from multiple situations (Fig. 1):

(i) a change of the niche envelope (overlap ≠ 1) due to symmetric niche expansion or contraction (hereafter called ‘unfilling’ in the case of invasions, because it corresponds to a part of the native niche that was not filled) in climatic space, thus not shifting the niche centroid (Fig. 1a); a species may expand both to warmer and colder conditions in a way that the average temperature-related niche position remains stable as it is observed for common spotted knapweeed invading North America (*Centaurea stoebe*, Fig. S1a, suppl. mat.);

(ii) a change of the niche centroid with displacement of the niche envelope (Fig. 1b-d) due to niche unfilling (e.g. black cherry tree invading Europe, *Prunus serotina*, Fig. S1b) and/or expansion (e.g. desert false indigo invading Europe, *Amorpha fruticosa*, Fig. S1c) in the invaded range;

or (iii) a change of the niche centroid only, without niche expansion or unfilling, due to a change of the density of occurrences within the same niche envelop in climatic space (Fig. 1e). The latter case can result from changes in competition, limited dispersal or availability of environmental conditions in the exotic range that reduce the density of species occurrences in some part of the niche space (Soberón & Peterson, 2011), changing the position of the centroid with only a weak impact on the niche limits, as shown for pinweed invading North America (*Erodium cicutarium*, Fig. S1d). Thus, a shift of the niche centroid between the native and the exotic range (Fig. 1b-e) can provide a first indication that a niche change occurred, but it is not sufficient to interpret its exact nature. And, reciprocally, an absence of a shift of the niche centroid does not mean that no niche shift occurred.

New indices were thus required to decompose niche comparisons to reveal two distinct components of niche changes: expansion and unfilling (Box 3, Fig. S1) (Petitpierre *et al.*, 2012). Unfilling (U) most commonly corresponds to the proportion of the native niche non-overlapping with the exotic niche, and expansion (E) refers to the proportion of the exotic niche non-overlapping with the native niche. These indices, as just defined, measure changes that are relative to one of the ranges (native or exotic), but they can also be measured with regard to the entire species distribution, where native and exotic ranges are pooled (Box 3). The pooled versions of E and U (E_p and U_p in Box 3) thus inform us about the species niche dynamic at the global scale of the study, but convey less information about our ability to predict species invasions from the native range (Box 3). E and U (and equivalently E_p and U_p) are recently published indices (Petitpierre *et al.*, 2012) that can easily be calculated from the same two main approaches previously described (Broennimann *et al.*, 2012), but provide much more information than simple overlap or centroid changes. Studies that found overall niche conservatism for invaders relied consistently on such complete set of niche change metrics (Table S1). Later, we will refer to the whole set of niche change metrics as the COUE scheme (Box 3).

Box 3: Metrics to quantify and decompose niche changes – The COUE scheme

The niche space of an exotic species can be classified into three categories: niche space occurring only in the exotic range (i.e. expansion, ε), in both exotic and native range (i.e. stability, σ) and only in the native range (i.e. unfilling, υ). Niche comparisons can then be made at two levels: (i) relative to the entire niche of the species, pooled from the two ranges (pooled ranges approach); or (ii) relative to the native or exotic ranges separately (range-specific approach). The table below presents a unified terminology (COUE, an acronym based on its main components, centroid shift, overlap, unfilling and expansion, as defined below) for niche comparisons and related metrics of niche changes.

Niche change component	Absolute component	Metric	
		Pooled ranges	Range-specific
Centroid shift	--	C	--
Expansion	ε	E_p	E
Stability	σ	S_p (≈ O)	S_n / S_e
Unfilling	υ	U_p	U

Centroid shift measures the change in mean niche position (and thus mean intensity) in the pooled ranges space, and thus no range-specific counterpart exists here. At the pooled-range level, niche stability (S_p) measures the proportion of niche assessed from the pooled native and exotic occurrences (possibly transformed into ENM predictions or densities in the environmental space, Fig. 3) present in both native and exotic ranges. This is similar to the niche overlap (O) assessed through Schoener's D or Hellinger's I (see Warren *et al.*, 2008, Broennimann *et al.*, 2012). The non-overlapping parts of the two niches (1 - S_p) can then be decomposed into global ratios of expansion (E_p) and unfilling (U_p) based on the pooled ranges. Decomposing niche changes relative to the pooled species distribution informs about the magnitude of niche changes at the global scale (i.e. relative to the entire realized niche of the species), but may not be informative about niche changes specific to either exotic or native ranges (as used e.g. in Petitpierre *et al.*, 2012). For example, the exotic niche can be very small relative to the entire pooled niche but entirely located in environments different from the native niche, in which case E_p would be very small although the entire invaded niche is distinct from the native niche and would hardly be predictable from the native range data. It is however possible to quantify a ratio of expansion (E) and unfilling (U) relatively to the exotic or native niches only, i.e. at the range-specific level. These inform us about the relative importance of changes in each exotic and native niche. In turn, niche stability can be assessed from the perspective of native or invaded niches separately, depending on whether it complements the relative expansion or relative unfilling ratios (S_n = 1 - U; S_e = 1 – E, respectively).

Dealing with available and non-analog climates between ranges

The availability of climatic conditions in geographic space matters when quantifying niche changes between ranges. Due to the niche-biotope duality (i.e. the correspondence between environmental and geographic spaces; Box 2), some conditions common in the exotic range may be rare in the native range (or the converse; Box 4) so that, without correction, one may detect niche shifts (measured with centroid change or overlap of percentile envelops) only because these conditions are more or less available in one range than in the other (Soberón & Nakamura, 2009, Soberón & Peterson, 2011).

Accounting for environmental availability is thus necessary and has been done so far in two ways. First, niche change metrics can be corrected by the distribution of the available environment, either by comparing the overlap between native and exotic niches with the overlap between native and exotic ranges (Mandle *et al.*, 2010), or by transforming species densities in the environmental space into species "occupancies" (i.e. the ratio of density of species to the density of available environment (Broennimann *et al.*, 2012; see also Dormann *et al.*, 2010). Second, niche metrics can be calculated only within the most common environments shared between native and exotic ranges (say within the shared portion of the 75th percentiles encompassing the prevailing conditions in each range (Petitpierre *et al.*, 2012). Removal of rare climates is however likely to have a strong impact on the results (with either approach) when the two ranges show important differences in climate availability. In this case we advise comparing analyses across a range of percentiles (say 75, 80, 85, 90, 95 and 100%) in order to see how the quantification of niche change can be affected by various levels of trimming (see suppl. online mat. in Petitpierre *et al.*, 2012) and to understand the implications (specific to each case study) for the interpretation of niche changes.

Box 4: The available climate and the analog/non-analog issue

The available environment is a subset of all possible environmental combinations (Box 2). The existence of non-available environments constrains niche shape and size (Jackson & Overpeck, 2000). For instance, places with very warm summer temperature (say >40°C) and very cold winter temperature (say <-20°C) do not currently exist on Earth (see Figure 1 in Jackson & Overpeck, 2000). When comparing the available environment in two areas, some habitats in one area (or time period) may be much more frequent or rare than in the other area (or time period), or some specific conditions found in one range may be totally absent from the other range. For instance, some very dry conditions of Western North America are not found in Western Europe (Broennimann *et al.*, 2007) and tropical conditions of the Tertiary in Europe are not observed anymore (Willis & McElwain, 2002). Conditions similar in two ranges or two time periods are called 'analog' and those differing 'non-analog' (Williams & Jackson, 2007, Fitzpatrick & Hargrove, 2009, or 'non-overlapping backgrounds' in Soberón & Peterson, 2011). Non-analog environments in an invaded range, or in the future, typically represent situations outside the range of values considered to quantify the native niche and not experienced by the species before invasion, and therefore lead to difficulty in interpreting niche shifts (Petitpierre *et al.*, 2012) and predicting species distributions (Fitzpatrick & Hargrove, 2009). Tools are available to define areas in the exotic range with climates analog to the native range. The simplest approach is to define a bounding box that encloses all the conditions present in the native range (e.g. BIOCLIM; Busby 1991). Any pixel in the exotic range outside of the bounding box range can be considered non-analogous. A more refined approach is the MESS analysis (Multivariate Environmental Similarity Surfaces, Elith *et al.*, 2010), an index of similarity reporting the closeness of a point described by a set of climate attributes (e.g. a pixel in the exotic range) to the distribution of these attributes within a population of reference points (e.g. the native range).

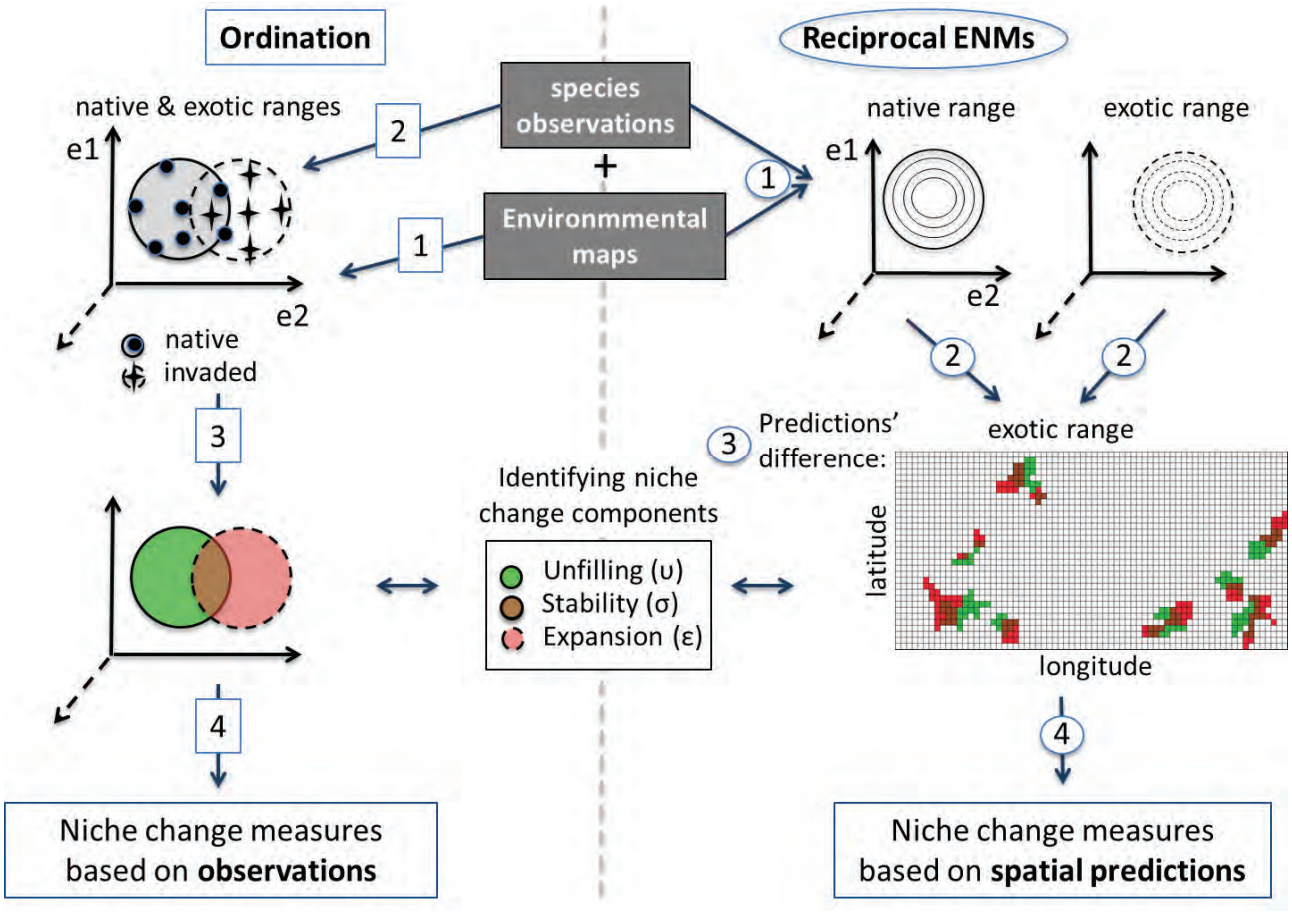


Figure 3: The two approaches commonly used to quantify niche changes between ranges (Box 1). Ordination is based on the observations directly, whereas ENM is based on predictions (see Broennimann *et al.*, 2012 and Box 1). Steps for ordination are (square numbers): 1. Definition of the reduced multidimensional environmental space; 2. Plot of the observations from each range in this space; 3. Comparison of the niche defined from observations in each range; 4. Calculation of the niche change metrics (see Box 3). Steps for ENMs are: 1. Fit of ENMs by relating field observations to environmental variables; 2. Projections of the ENMs in geographic space; 3. Compute difference in the projections; 4. Calculation of the niche change metrics (see Box 3). See main text for discussion of the respective strengths and weaknesses of the two approaches.

An extreme case of climate non-availability is when climate conditions exist only in one of the two ranges (Box 4, Williams & Jackson, 2007). These non-analog climates represent a severe problem when calculating niche change metrics, because no insight on the biology of the species in these non-analog climates can be learned from a comparison between ranges. This is because colonization of portions of environmental space not present in the native range cannot be considered unambiguously as resulting from niche evolution in the exotic range, and the interpretation of these situations thus remains speculative (Mandle *et al.*, 2010, Soberón & Peterson, 2011). A scientifically more rigorous approach to assess niche expansion, therefore, is to restrict the analyses to the shared, analog climatic conditions between the native and exotic ranges (e.g. Petitpierre *et al.*, 2012), and to provide measures of expansion in non-analog situations separately (Guisan *et al.*, 2012).

Studies that restricted their analyses to analog environments found niche conservatism to be dominant among invader species (Table S1). Complementary experimental approaches would then be needed to determine whether, for instance, expansions in non-analog conditions may represent a change of the fundamental niche (Pearman *et al.*, 2008*a*). This issue is particularly important because non-analog climates not only occur across space but will also occur over time due to climate change (Williams & Jackson, 2007). This is also the reason why projections of ecological models in non-analogous climates are considered unreliable (Fitzpatrick & Hargrove, 2009, Mandle *et al.*, 2010). Still, colonization of non-analog climates in the exotic range may represent relevant situations to consider from a management perspective, calling for separate ENM projections in both analog and non-analog climates in the invaded range (through fitting ENMs with pooled data from the native and exotic range, Broennimann & Guisan, 2008).

WHAT OTHER FACTORS COULD AFFECT THE QUANTIFICATION OF CLIMATIC NICHE CHANGES?

Range unfilling in the native range

Geographic range unfilling (not to be confused with niche unfilling) – i.e. when a species does not occupy all the geographic locations that have suitable conditions within its climatic niche – can occur in the native range as a result of non-equilibrium situations, such as ongoing post-glacial recolonization during the Holocene (Normand *et al.*, 2011), and can potentially affect the quantification of niche change. This problem is also known as the 'accessible area' issue (Barve *et al.*, 2011). But geographic range unfilling does not necessarily lead to niche unfilling in environmental space. For instance, it was shown that range unfilling can nonetheless translate into complete climatic niche filling for some tree species in Europe (Randin *et al.*, 2013). Range unfilling particularly affects niche quantification if the climates present in the unfilled geographic space are rare and/or not well represented – or even absent – in other parts of the range. However, published analyses generally calculate range filling based on a geographic projection of the realized niche at the time of the study (e.g. (Normand *et al.*, 2011)), and thus these documented cases of range unfilling cannot translate into niche unfilling. But range unfilling measured in other ways - e.g. field common garden experiments located beyond a species' current geographic and climatic range (Alexander *et al.*, 2012, Alexander, 2013, Hargreaves *et al.*, 2014) - may reveal niche unfilling.

Biased or incomplete sampling of species distributions

Another issue relates to the type and quality of species distribution data. Although it is important to cover an entire species' niche to assess niche change without bias, its complete native and exotic distribution ranges need not necessarily be considered. Because of the niche-biotope duality (Colwell & Rangel, 2009, Soberón & Nakamura,

2009, Box 2), the climatic niche of a species might well be fully captured even if only a part of its geographical distribution is sampled. However, and similarly to the issue of range unfilling, when geographic truncation leads to environmental truncation (Raes, 2012), niche change studies based on both ordinations and ENMs (Fig. 3) should be considered with care, because their conclusions will only be applicable to the climate space investigated and within analog climatic combinations between the two ranges. In these situations, approaches based on ecological niche models (ENMs, Guisan & Thuiller, 2005, Peterson *et al.*, 2011, Fig. 3) may be less reliable for spatial predictions, as they rely heavily on fitted species-environment response curves that could be biased (Thuiller *et al.*, 2004, Raes, 2012). In addition to environmental truncation, bias or errors in the geographic sampling of the distribution of a species may also bias measures of niche change. For instance, coarse atlas distribution data may portray a species in areas where it does not exist, while occurrence data (e.g. from herbaria) may under-represent or omit areas where the species occurs, both possibly affecting niche quantification.

Beyond macroclimate: microclimate and non-climatic factors

Climate is often seen as the main factor driving species distributions at large scales (Guisan & Thuiller, 2005), and most global-scale studies of niche changes in native (Pearman *et al.*, 2008*b*, Crisp *et al.*, 2009, Maiorano *et al.*, 2012) and exotic species (Broennimann *et al.*, 2007, Gallagher *et al.*, 2010, Medley, 2010; Lauzeral *et al.*, 2011, Petitpierre *et al.*, 2012) looked at changes in macroclimate (i.e. the coarse and large-scale climate that usually determines biomes). This primary role of macroclimate does not prevent finer climatic characteristics or other abiotic factors from affecting species distributions, such as the restriction to specialized habitats (e.g. mountain microclimates, stream banks or particular soil types) that must generally be characterized at a finer spatial grain (e.g. 1 km x 1 km) than that typically used in macroclimatic studies. Niche changes may be particularly observed in non-climatic components (such as soils) of a species' niche. For instance, Bertrand *et al.* (2012) showed that a shift of the climatic niche centroid can be observed when soil variables are included in the analyses. When shaping the distribution in the native range, these micro-scale factors could thus result in the detection of apparent macroclimatic niche expansion in the exotic range for two reasons: (i) part of the native macroclimate might not be occupied by the species due to spatial correlation with factors that hinder its occurrence (Bertrand *et al.*, 2012); or (ii) a species might occur under conditions in the native range that, within the coarse cells of macro-climatic maps, are scattered and marginal (and thus smoothed and hindered in niche analyses based on mean values within coarse cells), but are dominant in the exotic range and thus only revealed there in the niche quantification, causing an apparent niche shift.

However, these factors will only modify measures of macroclimatic niche change if: (i) their geographic distribution matches a restricted portion of the climatic niche in the native or exotic range, and (ii) this restricted portion is the one that shows niche change. In this regard, studies at a finer resolution (e.g. microclimate) and/or including non-climatic factors

would be useful for a more detailed understanding of niche dynamics in invaded ranges. But to be complementary to the strict macroclimatic niche studies conducted so far, findings based on macroclimate alone should be presented and compared to findings when microclimatic and non-climatic components are added (as for analog/non-analog climates), so that their relative effect can be properly assessed (e.g. Bertrand *et al.*, 2012).

TOWARD A UNIFYING FRAMEWORK: CONCLUSION AND REMAINING CHALLENGES

There has recently been a great diversity of studies examining climate niche change in exotic species (Table S1), some reporting dramatic niche changes (Broennimann *et al.*, 2007, Fitzpatrick *et al.*, 2007, Gallagher *et al.*, 2010, Medley, 2010). However, how many shifts occur in analog versus in non-analog climates, and whether these only occur in specific taxonomic groups or habitats, remains to be investigated. Among 36 studies including ca. 180 species, about 50% of the species showed overall a niche shift, with a higher prevalence among plants than animals, and a majority of the studies reporting niche shifts included only one or a few species (Table S1). It might therefore be that studies reporting a shift (rather than no shift) were preferentially published, especially considering that the only two studies that concluded overall niche conservatism among a large number of invader species used an ordination approach, relied on the most complete set of niche change metrics, and accounted for environmental availability (Table S1). Therefore, conclusions on niche shifts likely depend strongly on the organisms, methods and data used, and generalization about the frequency and drivers of niche shifts can only be based on a standardized and rigorous approach for quantifying niche shifts within each group. This could ultimately allow concluding if there are identifiable trends among niche shifts, or if niche changes are very idiosyncratic (i.e. species specific). In order to promote such standardization in future studies, we recommend:

- Using at least ordination, rather than only ENM, approaches to quantify climatic niche changes (see Broennimann *et al.*, 2012);

–Using as much as possible, within a same taxonomic group, the same set of variables used in previous studies on the same group, so that proper comparisons can be ensured; this does not prevent additionally testing niche changes with other sets of variables, if thought to be more meaningful to picture species' niches in the group considered;

- Disentangling all possible situations of niche change through measures of niche unfilling and expansion in complement to centroid shift and overlap metrics, at the two possible analytical levels (COUE scheme; Box 3);

- Correcting these niche change metrics to account for the density of occurrences and the available environment in both ranges (or time periods);

- Assessing whether niche metrics change when excluding rare climates along a range of percentiles, and when considering analog and non-analog environments separately; this

will ensure retaining all the necessary information for further interpretation and comparison of results from different studies.

We suggest three important remaining challenges for studies of realized niche changes during biological invasions:

1) Assessing climatic niche changes at finer scales and in combination with other non-climatic factors, such as differences in soils (Bertrand *et al.*, 2012), biota, and disturbances between the native and exotic range. High-resolution data are becoming increasingly available and standardized to be comparable across large spatial areas. They constitute avenues to provide complementary answers to questions on macroclimate niche changes, and to improve our ability to predict and anticipate invasions.

2) Assessing invasions in non-analog environments has been poorly addressed so far. As these situations cannot be predicted from the native range with static approaches, and thus their interpretations remain speculative, they require mechanistic approaches (e.g. Kearney & Porter, 2009) or experiments (see below). It is however a promising field of investigation that may deliver invaluable insights on colonization processes in non-analog situations while also improving assessments of biodiversity under future climate changes (Williams & Jackson, 2007). Retrospective studies that examine the details of invasion success and failure into particular non-analog climates, relative to the native climatic niche, could inform us of possible predictors of invasion into non-analog climates (e.g. for niche-based spatial predictions, Guisan *et al.*, 2012).

3) Although correlative niche shift studies of exotic species may guide experimental studies (Kueffer *et al.*, 2013), a dual approach has been rare so far (but see Hill *et al.*, 2013). Experimental studies on populations found in geographic areas where niche expansion occurred in the exotic range are needed to rigorously identify the related ecological or evolutionary causes, e.g. through rapid evolution (Sax *et al.*, 2007, Alexander, 2013), increased phenotypic plasticity (Hahn *et al.*, 2012) or biotic interactions (e.g. enemy-release, Alexander & Edwards, 2010). Similarly, information about unfilling can help identify interesting model systems (Kueffer *et al.*, 2013) for studying why some habitats and landscapes are more resistant to invasions, e.g. due to dispersal limitations (Barve *et al.*, 2011) or abiotic or biotic resistance (Richardson, 2011).

We expect that systematic use of this framework will substantially advance generalization about niche change, not only in invasion studies (including pests and diseases) but also in studies of niche conservatism between disjoint distributions (e.g. artic-alpine, Pellissier *et al.*, 2013) or across time in response to global change (Maiorano *et al.*, 2012).

ACKNOWLEDGEMENTS

We would like to acknowledge useful comments by Christophe Randin, Anibal Pauchard and Heinz Müller-Schärer. AG and CK led the study and all co-authors contributed to the review writing. AG, OB, and BP developed the figures. BP, CK and CD prepared table S1

(suppl. mat.). AG, OB and BP received support at the earlier stages of the study from the Swiss National Center of Competences in Research ‘Plant Survival’ (granted by the Swiss National Science Foundation), CK received partly support from USDA NRI grant no. 2006-35320-17360 to CD.

Glosary box

Analog climate: A combination of climate factors found in one area or time period that is within the envelop of climatic conditions found in a different area or time period used for comparison Williams and Jackson, 2007. Contrary: ‘non-analog climate’.

Accessible range: The geographic locations within a given area that are accessible to a species given its current distribution and the timescale considered in the study. It is thus conditional upon spatial configuration and the species’ dispersal ability Barve *et al.*, 2011, Soberon, 2007.

Available environment: the set of environmental conditions that exist in a given area Jackson and Overpeck, 2000 (Box 3). Synonyms: ‘realized environment’ (whole range, not species-specific), ‘background environment’.

Ecological niche model (ENM; also called species distribution or habitat suitability models): multivariate models fitting the niche of species by relating distribution observations with environmental variables measured at the same sites, and projected over a whole study area (see Guisan and Thuiller, 2005, Peterson *et al.*, 2011).

Exotic niche: The niche measured based on a species’ distribution in the exotic range. Synonyms: ‘naturalized niche’, ‘adventive niche’, ‘invaded niche’ or ‘invasive niche’ (for invasive species).

Exotic range: The geographic range where a species is not native. Synonyms: ‘naturalized range’, ‘adventive range’, ‘invaded range’ (for invasive species)

Exotic species: A species present in a region where it is not native, mostly due to human actions that enabled it to overcome biogeographical barriers Richardson, 2011. Synonyms: ‘alien species’, ‘non-native species’, ‘non-indigenous species’, ‘introduced species’.

Fundamental niche: The envelope of environmental (abiotic) conditions allowing populations to sustain themselves in an n-dimensional environmental space. It depicts the eco-physiological requirements of species Soberon, 2007. Synonyms: ‘Physiological niche’.

Native niche: The niche measured in the native range.

Native range: The complete geographic area where an exotic species is native.

Niche-biotope duality: The reciprocal correspondence between the niche conditions in multidimensional environmental space and the physical locations that a species actually occupies in geographical space (derived from Colwell and Rangel, 2009).

Niche centroid: the mean niche position in n-dimensional environmental space.

Niche conservatism: The tendency for species to retain their niche in space and time. Synonyms: ‘niche stability’.

Niche envelope: The envelope of conditions in multivariate environmental space defining a species’ niche. The boundary of the envelope can be defined in many different ways (e.g. percentiles; see Broennimann *et al.*, 2012).

Niche expansion: Proportion of the exotic niche non-overlapping with the native niche.

Niche overlap: the intersection of two niches in n-dimensional environmental space.

Niche shift: A change in the centroid (see above) or limits of the niche envelop in environmental space. Synonyms: niche change.

Niche stability: Proportion of the exotic niche overlapping with the native niche

Niche unfilling: Proportion of the native niche non-overlapping with the exotic niche.

Non-analog climate: See 'analog' climate.

Ordination: statistical approach used to represent the arrangement of a series of objects described by multiple descriptor variables into a reduced multidimensional space which axes represent combinations of the initial variables (see PCA).

PCA: Principal component analysis, a classical ordination approach (see above).

Potential niche: The intersection between the fundamental niche and the realized environment (see Soberón and Nakamura, 2009, Jackson and Overpeck, 2000).

Rare climate: Climatic conditions poorly represented overall within an area during a given time period.

Realized niche: The environmental (abiotic) niche of a species as quantified from field observations, i.e. the fundamental niche modulated by biotic exclusions, population dynamics (such as source-sink dynamics) and dispersal limitations Colwell and Rangel, 2009, Soberon, 2007. Synonyms: 'Ecological niche'.

Schoener's D: The most common measure of niche overlap (see Broennimann *et al.*, 2012, Warren *et al.*, 2008).

REFERENCES

Ackerly, D.D. (2003) Community assembly, niche conservatism, and adaptive evolution in changing environments. *International Journal of Plant Sciences*, **164**, S165-S184.

Alexander, J.M. (2013) Evolution under changing climates: climatic niche stasis despite rapid evolution in a non-native plant. *Proceedings of the Royal Society B*, **280**, 1446.

Alexander, J.M. & Edwards, P.J. (2010) Limits to the niche and range margins of alien species. *Oikos*, **119**, 1377-1386.

Alexander, J.M., van Kleunen, M., Ghezzi, R. & Edwards, P.J. (2012) Different genetic clines in response to temperature across the native and introduced ranges of a global plant invader. *Journal of Ecology*, **100**, 771-781.

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

Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810-1819.

Bertrand, R., Perez, V. & Gegout, J.C. (2012) Disregarding the edaphic dimension in species distribution models leads to the omission of crucial spatial information under climate change: the case of Quercus pubescens in France. *Global Change Biology*, **18**, 2648-2660.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, **4**, 585-589.

Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701-709.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481-497.

Colwell, R.K. & Rangel, T.F. (2009) Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19651-19658.

Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P. & Linder, H.P. (2009) Phylogenetic biome conservatism on a global scale. *Nature*, **458**, 754-756.

Dormann, C.F., Gruber, B., Winter, M. & Hermann, D. (2010) Evolution of climatic niches in European mammals? *Biology Letters*, **6**, 229-232.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330-342.

Elton, C.S. (1958) *The ecology of invasions by animals and plants*. Methuen, London.

Fitzpatrick, M.C. & Hargrove, W.W. (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation*, **18**, 2255-2261.

Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, **16**, 24-33.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology*, **98**, 790-799.

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

Guisan, A., Petitpierre, B., Broennimann, O., Kueffer, C., Randin, C. & Daehler, C. (2012) Response to comment on "Climatic niche shifts are rare among terrestrial plant invaders". *Science*, **338**, 193.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16**, 1424-1435.

Hahn, M.A., van Kleunen, M. & Müller-Schärer, H. (2012) Increased phenotypic plasticity to climate may have boosted the invasion success of polyploid *Centaurea stoebe*. *PLoS ONE*, **7**, e50284.

Hargreaves, A.L., Samis, K.E. & Eckert, C.G. (2014) Are Species' Range Limits Simply Niche Limits Writ Large? A Review of Transplant Experiments beyond the Range. *The American Naturalist*, **183**, 157-173.

Hill, M.P., Chown, S.L. & Hoffmann, A.A. (2013) A predicted niche shift corresponds with increased thermal resistance in an invasive mite, *Halotydeus destructor*. *Global Ecology & Biogeography*, **22**, 942-951.

Hutchinson, G.E. (1957) Population Studies - Animal Ecology and Demography - Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415-427.

Jackson, S.T. & Overpeck, J.T. (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, **26**, 194-220.

Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334-350.

Kueffer, C., Pyšek, P. & Richardson, D.M. (2013) Integrative invasion science: model systems, multi-site studies, focused meta-analysis, and invasion syndromes. *New Phytologist*, **200**, 615-633.

Lauzeral, C., Leprieur, F., Beauchard, O., Duron, Q., Oberdorff, T. & Brosse, S. (2011) Identifying climatic niche shifts using coarse-grained occurrence data: a test with non-native freshwater fish. *Global Ecology and Biogeography*, **20**, 407-414.

Maiorano, L., Cheddadi, R., Zimmerman, N.E., Pellissier, L., Petitpierre, B., Pottier, J., Laborde, H., Hurdu, B.I., Pearman, P.B., Psomas, A., Singarayer, J.S., Broennimann, O., Vittoz, P., Dubuis, A., Edwards, M.E., Binney, H.A. & Guisan, A. (2012) Building the niche through time: using 13,000 years of data to predict the effects of climate change on three tree species in Europe. *Global Ecology & Biogeography*, **22,** 302-317.

Mandle, L., Warren, D.L., Hoffmann, M.H., Peterson, A.T., Schmitt, J. & von Wettberg, E.J. (2010) Conclusions about Niche Expansion in Introduced *Impatiens walleriana* Populations Depend on Method of Analysis. *Plos One*, **5**

Medley, K.A. (2010) Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus Skuse* (*Culicidae*), revealed by reciprocal distribution models. *Global Ecology and Biogeography*, **19**, 122-133.

Normand, S., Ricklefs, R.E., Skov, F., Bladt, J., Tackenberg, O. & Svenning, J.C. (2011) Postglacial migration supplements climate in determining plant species ranges in Europe. *Proceedings of the Royal Society B-Biological Sciences*, **278**, 3644-3653.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008*a*) Niche dynamics in space and time. *Trends in Ecology & Evolution*, **23**, 149-158.

Pearman, P.B., Randin, C.F., Broennimann, O., Vittoz, P., van der Knaap, W.O., Engler, R., Le Lay, G., Zimmermann, N.E. & Guisan, A. (2008*b*) Prediction of plant species distributions across six millennia. *Ecology Letters*, **11**, 357-369.

Pellissier, L., Bråthen, K.A., Vittoz, P., Yoccoz, N.G., Dubuis, A., Meier, E.S., Zimmermann, N.E., Randin, C.F., Thuiller, W., Garraud, L., Van Es, J. & Guisan, A. (2013) Thermal niches are more conserved at cold than warm limits in arctic-alpine plant species. *Global Ecology and Biogeography*, **22**, 933-941.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, **38**, 817-827.

Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with Solenopsis invicta and *Solenopsis richteri*. *Global Ecology and Biogeography*, **17**, 135-144.

Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martinez-Meyer, E., Nakamura, M. & Araujo, M.B. (2011) *Ecological niches and geographic distributions*. Princeton University Press, Princeton, USA.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science*, **335**, 1344-1348.

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

Raes, N. (2012) Partial versus Full Species Distribution Models. *Natureza & Conservação*, **10**, 127-138.

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

Randin, C.F., Paulsen, J., Vitasse, Y., Kollas, C., Wohlgemuth, T., Zimmermann, N.E. & Korner, C. (2013) Do the elevational limits of deciduous tree species match their thermal latitudinal limits? *Global Ecology and Biogeography*, **22**, 913-923.

Richardson, D.M. (2011) *Fifty years of invasion ecology: the legacy of Charles Elton*. Blackwell Publishing, Oxford.

Richardson, D.M. & Pysek, P. (2008) Fifty years of invasion ecology - the legacy of Charles Elton. *Diversity and Distributions*, **14**, 161-168.

Rödder, D. & Lötters, S. (2009) Niche shift versus niche conservatism? Climatic characteristics of the native and invasive ranges of the Mediterranean house gecko (*Hemidactylus turcicus*). *Global Ecology and Biogeography*, **18**, 674-687.

Sax, D.F., Early, R. & Bellemare, J. (2013) Niche syndromes, species extinction risks, and management under climate change. *Trends in Ecology & Evolution*, **28**, 517-523.

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S.D., Grosberg, R.K., HastingS, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I. & Rice, W.R. (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution*, **22**, 465-471.

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

Soberón, J. & Peterson, A.T. (2011) Ecological niche shifts and environmental space anisotropy: a cautionary note. *Revista Mexicana de Biodiversidad*, **82**, 1348-1355.

Soberón, J. & Nakamura, M. (2009) Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19644–19650.

Strubbe, D., Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography*, **22**, 962-970.

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

Venette, R.C., Kriticos, D.J., Magarey, R.D., Koch, F.H., Baker, R.H.A., Worner, S.P., Raboteaux, N.N.G., McKenney, D.W., Dobesberger, E.J., Yemshanov, D., De Barro, P.J., Hutchison, W.D., Fowler, G., Kalaris, T.M. & Pedlar, J. (2010) Pest risk maps for invasive alien species: a roadmap for improvement. *Bioscience*, **60**, 349-362.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. *Evolution*, **62**, 2868-2883.

Warren, D.L., Glor, R.E. & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607-611.

Webber, B.L., Le Maître, D.C. & Kriticos, D.J. (2012) Comments on "Climatic niche shifts are rare among terrestrial plant invaders". *Science*, **338**, 193.

Williams, J.W. & Jackson, S.T. (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, **5**, 475-482.

Willis, K.J. & McElwain, J.C. (2002) *The evolution of plants*. Oxford University Press.

SUPPLEMENTARY MATERIAL

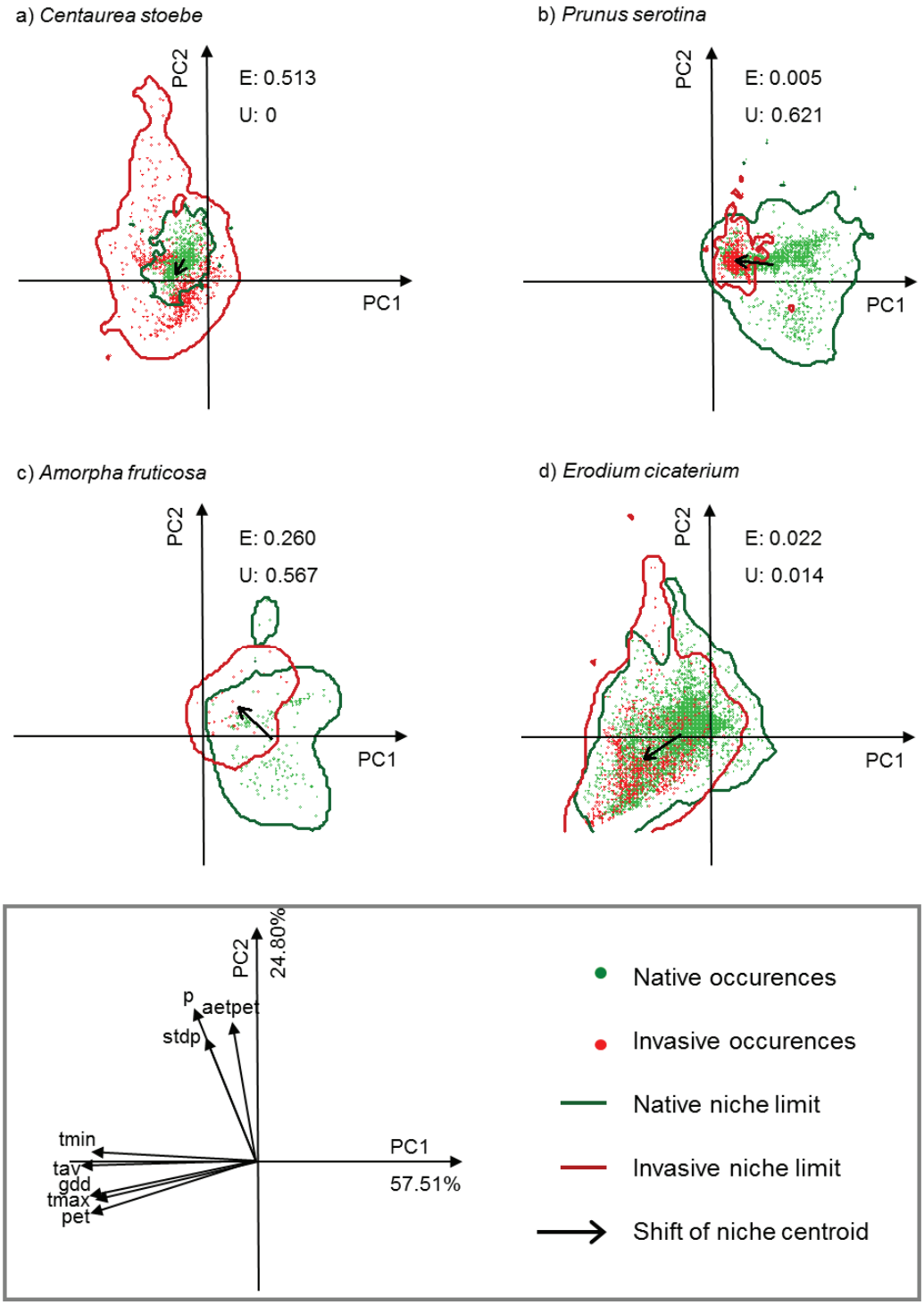
Table S1: Review of 36 studies comparing the niche of exotic species between their native and exotic ranges. We did not consider studies based on ENM prediction accuracy only, as these emphasize model transferability. 'Approach': type of data used for niche comparison, either species observations along individual environmental variables (Univariate), observations in multiple environmental space (Ordination), or ecological niche model predictions (ENM pred). 'Metric of niche change': metric(s) used for niche comparison, either shift of centroid (C), niche overlap (O) and/or one of the more dynamic components of niche change that are niche unfilling and/or expansion (U/E). * metrics statistically tested. 'Climatic availability': indicates whether differences in climatic availability between native and exotic ranges were taken into account. Parentheses mean that metrics or climatic availability were only considered qualitatively and not rigorously quantified. 'Reported niche shifts': synthesizes if the study concludes that there was a niche shift, with the number of species given in parentheses. 'References': paper belonging to each category of approach, with indication of the type of organisms studied between square brackets. Among the 36 studies, reporting on ca. 180 case studies at the species level, a niche shift was reported in slightly more than 50% of cases (ca 21% plants, 29% animals), no shift in slightly more than 42% of cases (12% plants, 30% animals), and no conclusion in 8% of cases. Thus there were slightly more animals than plants showing shifts.. But interestingly, among these studies, the prevalence of detected climate niche shift was higher among plants (40/63 case studies; Table S1) than among animals (54/110; Table S1). Interestingly, the two studies that concluded majoritarily on conservatism over a large set of species were the ones relying on the complete set of niche change metrics (O*/U/E) and accounting for climatic availability. This contrasting evidence may, therefore, result from different types of niche changes, biological and/or methodological study contexts, data types, species characteristics or methods being used (Peterson & Nakazawa, 2008, Mandle *et al.*, 2010, Peterson, 2011, Soberón & Peterson, 2011, Broennimann *et al.*, 2012) with confounding effects. Without methodological standardization, this makes comparative analyses and sound interpretation of the possible mechanisms behind the reported niche dynamics difficult.

Approach	**Metric of niche change**	**Climatic availability**	**reported niche shifts (# species)**	**References [taxonomic group]**
Univariate	C*	yes	yes(3)	Kolbe, J. J. *et al.* 2012. *Ecol. Evol*. [Vertebrate]
	(O)	no	yes(1)	Rödder, D. *et al.* 2009. *PLoS ONE* [Vertebrate]
	(U/E)	no	yes(18)	Lauzeral, C. *et al.* 2011. *Glob. Ecol. Biogeogr*. [Aquatic Vertebrate]
Univariate + Ordination	C*/O	no	yes(1)	Gallardo, B. *et al.* 2013. *J. Biogeograph*. [Aquatic Invertebrate]
Ordination	C*	no	yes(1)	Barbosa, F. G. *et al.* 2013. *Austral Ecol*. [Plant]
			yes(1)	Da Mata, R. A. *et al.* 2010. *Biol. Invasions* [Invertebrate]
			yes(2)	Steiner, F. *et al.* 2008. *Divers. Distrib.* [Invertebrate]
		(yes)	yes(2), no(1)	Beaumont, L. J. *et al.* 2009. *Divers. Distrib*. [Plant]
			yes(1)	Broennimann, O. *et al.* 2007. *Ecol. Lett.* [Plant]

	(O)	(yes)	yes(4)	Capinha, C. *et al.* 2011. *Ecography* [Invertebrate]
		no	yes(1)	Fitzpatrick, M. C. *et al.* 2007. *Glob. Ecol. Biogeogr.* [Invertebrate]
			yes(1)	Mau-Crimmins, T. M. *et al.* 2006. *Ecol. Model.* [Plant]
	O/U/E	yes	no(1)	Alexander, J. M. 2013. *Proc. R. Soc. Lond. B* [Plant]
	O*	yes	yes(2)	Broennimann, O. *et al.* 2012. *Glob. Ecol. Biogeogr.* [Plant and Invertebrate]
			yes(1)	Di Febbraro, M. *et al.* 2013. *PLoS ONE* [Vertebrate]
	O*/U/E	yes	yes(8), no(20)	Strubbe, D. *et al.* 2013. *Glob. Ecol. Biogeogr*. [Vertebrate]
	C*/O*/U/E	yes	yes(7), no(43)	Petitpierre, B. *et al.* 2012. *Science* [Plant]
	C*/(U/E)	no	yes(20), no(6)	Gallagher, R. V. *et al.* 2010. *J. Ecol*. [Plant]
	C*/O	no	yes(3)	Larson, E. R.& J. D. Olden. 2012. *Glob. Ecol. Biogeogr*. [Aquatic Invertebrate]
ENM pred.	(O)	no	no(1)	Peterson, A. T. and Y. Nakazawa. 2008. *Glob. Ecol. Biogeogr*. [Invertebrate]
	O	no	yes (7), no (4)	Donaldson, J. E. *et al.* 2014. *Glob. Chang. Biol*. [Plants]
	O*	yes	no conclusion (1)	Angetter, L. S. *et al.* 2011. *Biol. J. Linn. Soc.* [Vertebrate]
			no conclusion (10)	Fernandez, M. *et al.* 2012. *Ecol. Model*. [Plant, Vertebrate and Invertebrate]
			no conclusion (1)	Rödder, D. & S. Lötters. 2009. *Glob. Ecol. Biogeogr*. [Vertebrate]
Univariate + ENM pred.	O*	no	yes (1)	Schulte U. *et al.* 2012. *Glob. Ecol. Biogeogr*. [Vertebrate]
	C/O*	yes	no conclusion (1)	Rödder, D. & S. Lotters. 2010. *Naturwissenschaften* [Vertebrate]
	C*/O*	yes	no conclusion (1)	Mandle, L. *et al.* 2010. *PLoS ONE* [Plant]
Univariate + Ordination + ENM pred.	O	yes	yes(1)	Zhu, G. *et al.* 2012. *PLoS ONE* [Invertebrate]
	C*/O*	no	no(1)	Guo, W. Y. 2013. *Glob. Chang. Biol*. [Plant]
Ordination + ENM pred.	C*/O	no	no(1)	Palaoro, A. V. 2013. *Freshwater Biol.* [Aquatic Vertebrate]

	C*/O*	yes	yes(1)	Larson, E. R. *et al.* 2010. *Ecosphere* [Invertebrate]
			yes(1)	Medley, K. A. 2010. *Glob. Ecol. Biogeogr*. [Invertebrate]
			yes(1)	Morehouse, R. L. & M. Tobler. 2013. J. *Crustac. Biol*. [Invertebrate]
			yes(1)	Mukherjee, A. *et al.* 2012. *Biol. Invasions*. [Plant]
			yes(2)	Petersen, M. J. 2013. *Biol Invasions* [Invertebrate]
	O*	yes	no(1)	Zhu, G. et al, 2013. *Biol. Invasions* [Invertebrate]

Figure S1: Different real case of niche changes during invasions from Petitpierre *et al.* (2012): native and invasive distribution (represented by green and red dots respectively) in the environmental space depicted by the two first axes (PC1 and PC2) of a principal component analysis summarizing 57.51% and 24.80% respectively of the variation in climatic conditions available in North America and Eurasia. Climatic variables used in Petitpierre *et al.* (2012) are ratio of actual and potential evapotranspiration (aetpet), annual precipitations (p), precipitation seasonality (stdp), minimum temperature of the coldest month (tmin), average annual temperature (tav), growing degree days (gdd), maximum temperature of the hottest month (tmax) and potential evapotranspiration (pet). Native and invasive niche limits after kernel smoothing Petitpierre *et al.*, 2012 are represented by green and red lines respectively whereas difference between native and invasive niche centroids (i.e. niche shifts) are represented by the arrow. Niche expansion (E) and unfilling (U) are provided for each species.



REFERENCES

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481-497.

Mandle, L., Warren, D.L., Hoffmann, M.H., Peterson, A.T., Schmitt, J. & von Wettberg, E.J. (2010) Conclusions about Niche Expansion in Introduced *Impatiens walleriana* Populations Depend on Method of Analysis. *Plos One*, **5**

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, **38**, 817-827.

Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, **17**, 135-144.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science*, **335**, 1344-1348.

Soberón, J. & Peterson, A.T. (2011) Ecological niche shifts and environmental space anisotropy: a cautionary note. *Revista Mexicana de Biodiversidad*, **82**, 1348-1355.

CHAPTER 1.2

Climatic niche shifts are rare among terrestrial plant invaders

Blaise Petitpierre[1], Christoph Kueffer [2,3], Olivier Broennimann[1], Christophe Randin[4], Curtis Daehler[2] & Antoine Guisan[1]

[1] Department of Ecology & Evolution, University of Lausanne, 1015 Lausanne, Switzerland

[2] Department of Botany, University of Hawaii at Manoa, 3190 Maile Way, Honolulu, HI 96822, USA

[3] Institute of Integrative Biology, ETH Zürich, Universitätstrasse 16, 8092 Zürich, Switzerland

[4] Institute of Integrative Biology, ETH Zürich, Universitätstrasse 16, 8092 Zürich, Switzerland

This chapter is published in *Science* (DOI: 10.1126/science.1215933)

My contribution to the project: I gathered the data, developed the framework to quantify niche shifts, ran all the analyses and wrote the initial draft of the paper.

ABSTRACT

The assumption that climatic niche requirements of invasive species are conserved between their native and invaded ranges is key to predicting the risk of invasion. However, this assumption has been challenged recently by evidence of niche shifts in some species. Here, we report the first large-scale test of niche conservatism for 50 terrestrial plant invaders between Eurasia, North America and Australia. We show that when analog climates are compared between regions, fewer than 15% of species have more than 10% of their invaded distribution outside their native climatic niche. These findings reveal that significant niche shifts are rare in terrestrial plant invaders, providing new support for an appropriate use of ecological niche models for the prediction of both biological invasions and responses to climate change.

REPORT

Niche conservatism in space and time is a key assumption for predicting the impact of global change on biodiversity (Pearman *et al.*, 2008, Wiens and Graham, 2005). It is particularly important for the anticipation of biological invasions, which can cause severe damage to biodiversity, economies and human health (Vilà *et al.*, 2009). Niche conservatism implies that species tend to grow and survive under the same environmental conditions in native and invaded ranges (Wiens and Graham, 2005). However, the generality of this assumption is challenged by recent evidence suggesting that the climatic niche occupied by species may not be conserved between their native and invaded ranges, as documented by observed niche shifts for plants (Broennimann *et al.*, 2007, Gallagher *et al.*, 2010), insects (Fitzpatrick *et al.*, 2007, Medley, 2010) and fishes (Lauzeral *et al.*, 2011). Yet, several of these studies have focused on a single species (e.g. Broennimann *et al.* 2007, Fitzpatrick *et al.* 2007, Medley 2010) or have used controversial niche overlap metrics (e.g. Gallagher *et al.*, 2010, Lauzeral *et al.*, 2011; based on 26 and 18 spp respectively), making it difficult to assess the generality of this phenomenon among alien invasive species. Therefore, the question of whether niche shifts represent a prominent or unusual phenomenon among alien invasive species remains largely unresolved (Alexander and Edwards, 2010).

There are two major reasons why niche conservatism during biological invasion needs further investigation. First, anticipation is the most effective management strategy (Leung *et al.*, 2002) and niche conservatism is a strong and necessary assumption to predict invasion risk for specific regions (Wiens and Graham, 2005, Pearman *et al.*, 2008). Ecological niche models (ENM, Guisan and Thuiller, 2005, Supp. Mat.), the most commonly used predictive tool in this regard, are traditionally calibrated using native species distributions and then projected onto other continents to highlight areas susceptible to invasions (Thuiller *et al.*, 2005). Second, detecting significant deviations from niche conservatism may highlight invasive species that are characterized by ecological (Klironomos, 2002, Hierro *et al.*, 2006) or evolutionary changes (Lavergne and

Molofsky, 2007, Xu *et al.*, 2009) during invasions, helping us understand when such changes are likely to occur, which is crucial in an era of rapid climate change.

When the niche of a species changes, its mean position (centroid) is likely to move within a multivariate environmental niche space. However, describing the shift of the centroid (Broennimann *et al.*, 2007, Gallagher *et al.*, 2010, Medley, 2010) falls short in helping to understand processes affecting the niche, because niche changes can affect both the position and the shape of a niche. This is for example, the case when species expand to new climates at one particular niche margin (Broennimann *et al.*, 2007, Pearman *et al.*, 2008) and only partially fill the niche (i.e. unfilling) at another (Welk, 2004) (e.g., due to dispersal limitation) (Fig. S1). Assuming a species is at equilibrium in its native range (i.e., has colonized all suitable environments), then expansion to climates that are new to the species but available in the native range should be considered unambiguously as niche shifts (Supp. Mat., Fig. S1), i.e., resulting from changes in biotic interactions or rapid evolution of the species (Pearman *et al.*, 2008). This conceptual approach to detecting niche shifts is important because analyses of empirical field data alone cannot determine whether the expansion to climates not available in the native range (i.e., non-analog climates) represents a true niche shift or the filling of a pre-adapted niche. On the other hand, unfilling in the invaded range is more likely due to dispersal limitation, because biological invasions are recent and ongoing phenomena.

Niche changes due to unfilling have been considered niche shifts in previous studies (Broennimann *et al.*, 2007, Fitzpatrick *et al.*, 2007, Gallagher *et al.*, 2010, Medley, 2010) but our analyses (see Supp. Mat) reveal that many of these reflect ongoing colonization instead, indicating that the species is likely to invade additional geographic regions in the future (Thuiller *et al.*, 2005). Thus, metrics of niche shift are very sensitive to the underlying statistical and conceptual assumptions and a solid conceptual foundation for identifying ecologically meaningful and statistically significant niche changes has only recently been developed (Supp. Mat., Warren *et al.*, 2008, Mandle *et al.*, 2010, Broennimann *et al.*, 2011).

Here, we disentangle and quantify the amount of niche overlap, niche expansion and niche unfilling (see Fig. S1 and S2) for 50 Holarctic terrestrial alien angiosperms (Tables S1 and S2). Plants are appropriate for this test because their distributions are largely limited by climatic factors (Woodward, 1987), a necessary condition to assess niche conservatism. Our sample includes many of the major plant invaders between North America (NA) and Eurasia (EU) and also many of the most anciently introduced EU species in NA. The reciprocal comparison of EU and NA invaders provides an important test of niche conservatism because it is the only pair of two large, separated landmasses with a largely overlapping climate space and a long history of reciprocal anthropogenic exchanges of floras (Crosby, 1986, Seipel *et al.*, 2011). When available, the distribution of these species in Australia (AU, Table S3), where none is native, was used to provide additional, independent insights into patterns of niche filling when climatic availability, although partly overlapping, is overall very different from the native range. Geographical distributions

(resolution = 0.5°, approximately 50 km) were projected onto climate space following a niche quantification framework correcting for species densities and climatic availability in both the native and invaded range (Supp. Mat., Broennimann *et al.*, 2011). This approach tests for niche conservatism and robustly quantifies the amount of niche overlap, expansion and unfilling in the invaded range.

We find little evidence of niche expansion associated with invasion of new regions. Our results for the Holarctic reveal that, although levels of niche overlap among species vary between 17% and 64% (Fig. 1, Table S5), niche conservatism is observed for 46% of species (23) between the native and invaded range in EU and NA (similarity test with a significance level ≤ 0.05; Fig. 1, Table S5). NA species show higher propensity toward niche similarity (13 out of 20 species). In contrast to comparisons between EU and NA, niche similarity tests for Australia are significant for all species (Table S6) despite more pronounced climatic differences between AU and both EU and NA, respectively, than between EU and NA. This indicates that in AU, Holarctic invasive species remain in Holarctic climates and are rarely found in new climates. In other words, when considering the available climate in the invaded range, species colonize climatic conditions close to the ones colonized in their native range.

Further differentiating non-overlap situations into cases of unfilling or expansion reveals that in the Holarctic only 14% of the studied species (7) show more than 10% expansion, with only one outlier species - spotted knapweed (*Centaurea stoebe*) - showing >50% expansion (Fig. 1, Fig. 2, Table S5). Previous studies also reported an important niche shift for this species (Broennimann *et al.*, 2007), possibly caused by evolutionary (Treier *et al.*, 2009) and/or ecological processes (Hierro *et al.*, 2006), notably through hybridization (Blair and Hufbauer, 2009, Broennimann *et al.*, 2007,) and enhanced competitive strength in the invaded range (Callaway *et al.*). Interestingly, there is also evidence of novel genetic admixing (repeated introductions or hybridization) and reduced impacts of competitors and enemies in many of the other studied species (e.g., Blair and Hufbauer, 2009, Durka *et al.*, 2005, Novak and Mack, 2001, Pairon *et al.*, 2010) but these species did not show any major niche expansion, indicating that these mechanisms do not necessarily lead to niche expansion. Indeed, niche unfilling is a more widespread phenomenon with 48% of species (24) showing more than 10% of their native niche unfilled in the invasive range (Fig. 1 and Fig. 3). Patterns in Australia confirm these Holarctic findings, i.e., niche expansion is uncommon compared to unfilling (Fig. 2, Fig. 4, Fig. S4, Table S6).

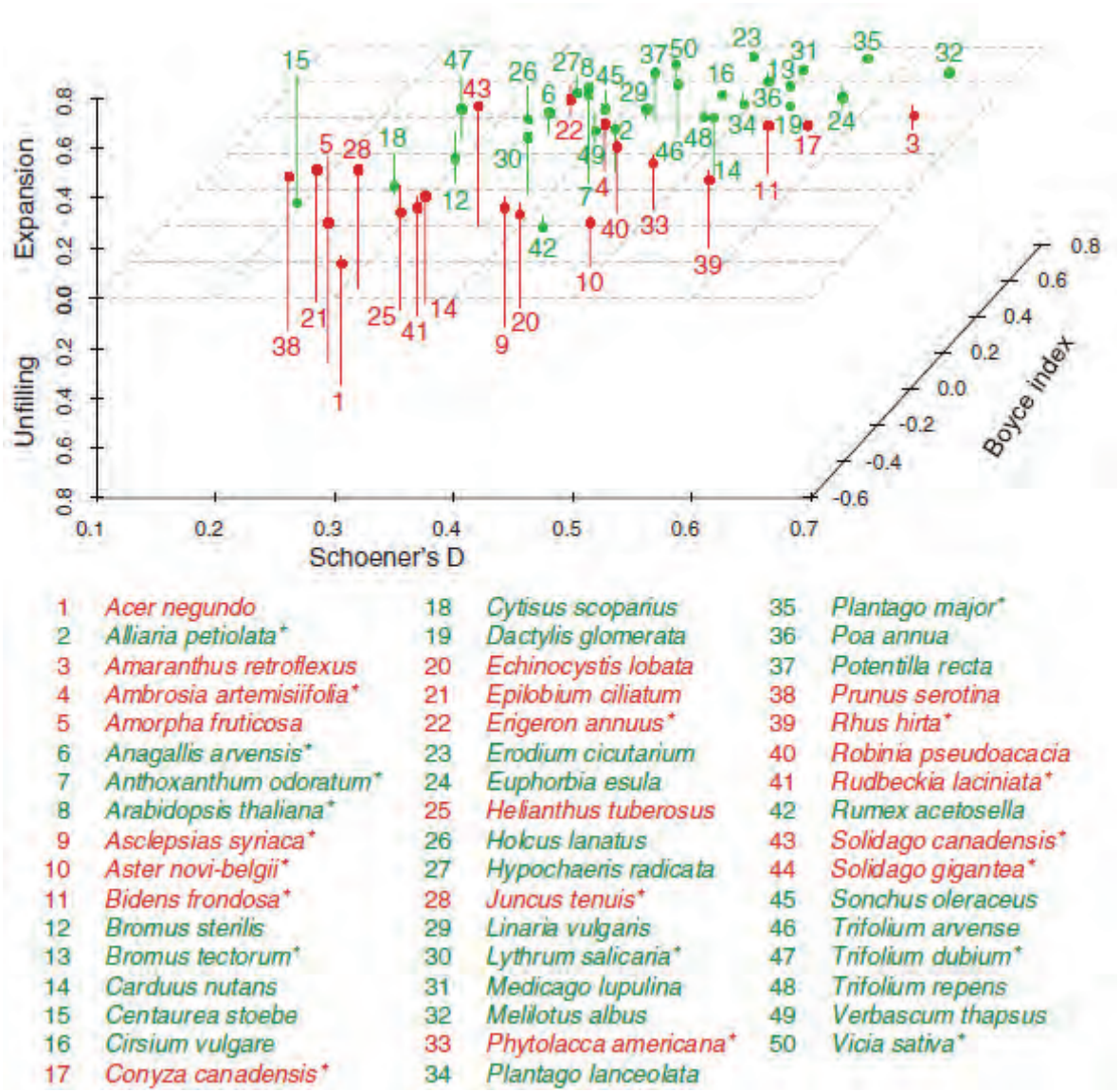


Figure 1: Niche changes between native and invaded ranges in Eurasia (EU) and North America (NA). Vertical segments represent the magnitude of niche changes for each species. Extensions above and below the zero plane indicate expansion and unfilling, respectively. Intersections with the zero plane are shown with dots. Green (EU) and red (NA) colors indicate species origin. Niche change indices are plotted over two niche overlap indices, Schoener's D and the Boyce index evaluation of ecological niche models (ENM) calibrated in the native range and projected onto analog climates in the invaded range. Stars show species with a significant niche overlap between native and invaded range based on a similarity test.

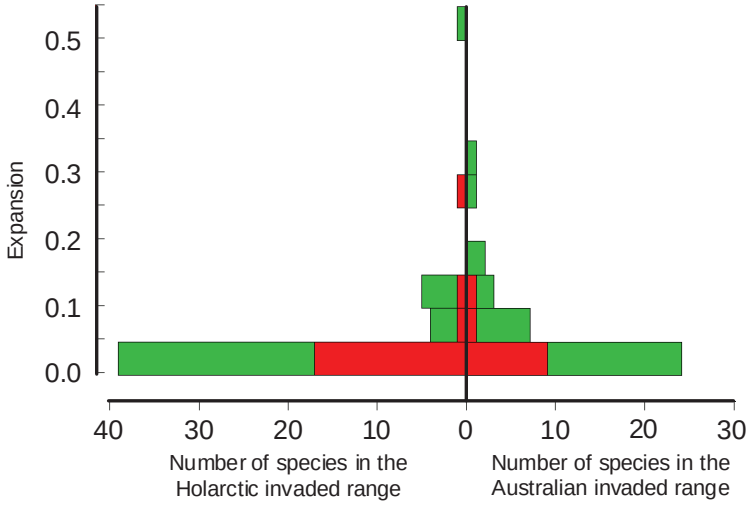


Figure 2: Expansion in Holarctic and Australian invaded ranges. The expansion index is analogous to the proportion of the invasive distribution in novel environments. NA and EU species origins are shown with red and green colors respectively

The biogeographical origin of the species provides further insights into niche comparisons between native and invaded ranges (Fig. 3 and Fig. 4). In general, EU species show less niche unfilling and more expansion in NA and AU than NA species in EU and AU, thus mirroring biogeographical patterns of invasibility, which show higher invasion rates of NA compared to EU (Seastedt and Pyšek, 2011). Differences in the geographic arrangement of EU versus NA could account for the difference in niche unfilling. In particular, climate varies on a shorter distance along latitudinal gradients in NA than EU and may allow more rapid expansion into different climates in NA (Rejmanek, 2000). However, this does not explain why EU species also show less niche unfilling in AU than NA species. Biome conservatism, frequent across long evolutionary time scales (Crisp *et al.*, 2009) and highly expected in the case of invasive species (Thuiller *et al.*, 2005), may further explain niche differences between areas differing in biome availability (Fig. 3 and Fig. 4). In NA and AU, EU species expansions occur toward hotter and drier niche limits, corresponding in NA to the median climatic conditions of temperate coniferous forests, which are mostly absent in EU (Fig. 3). The lower prevalence of niche unfilling in EU species may relate to the longer history of weed selection in human-disturbed landscapes in Europe and earlier colonization paths from Europe to other continents (Seipel *et al.*, 2011, Crosby, 1986). However, when testing the effect of minimum residence time on niche expansion, overlap, unfilling and total change magnitude, we found no significant effect (Table S5), suggesting that other drivers, such as human-mediated propagule pressure, likely prevail. Movement of human settlements was far more important from EU towards NA and AU than the opposite (Seastedt and Pyšek, 2011), as shown by higher numbers of Eurasian invaders worldwide (Seipel *et al.*, 2011) and this could explain less unfilling among EU species.

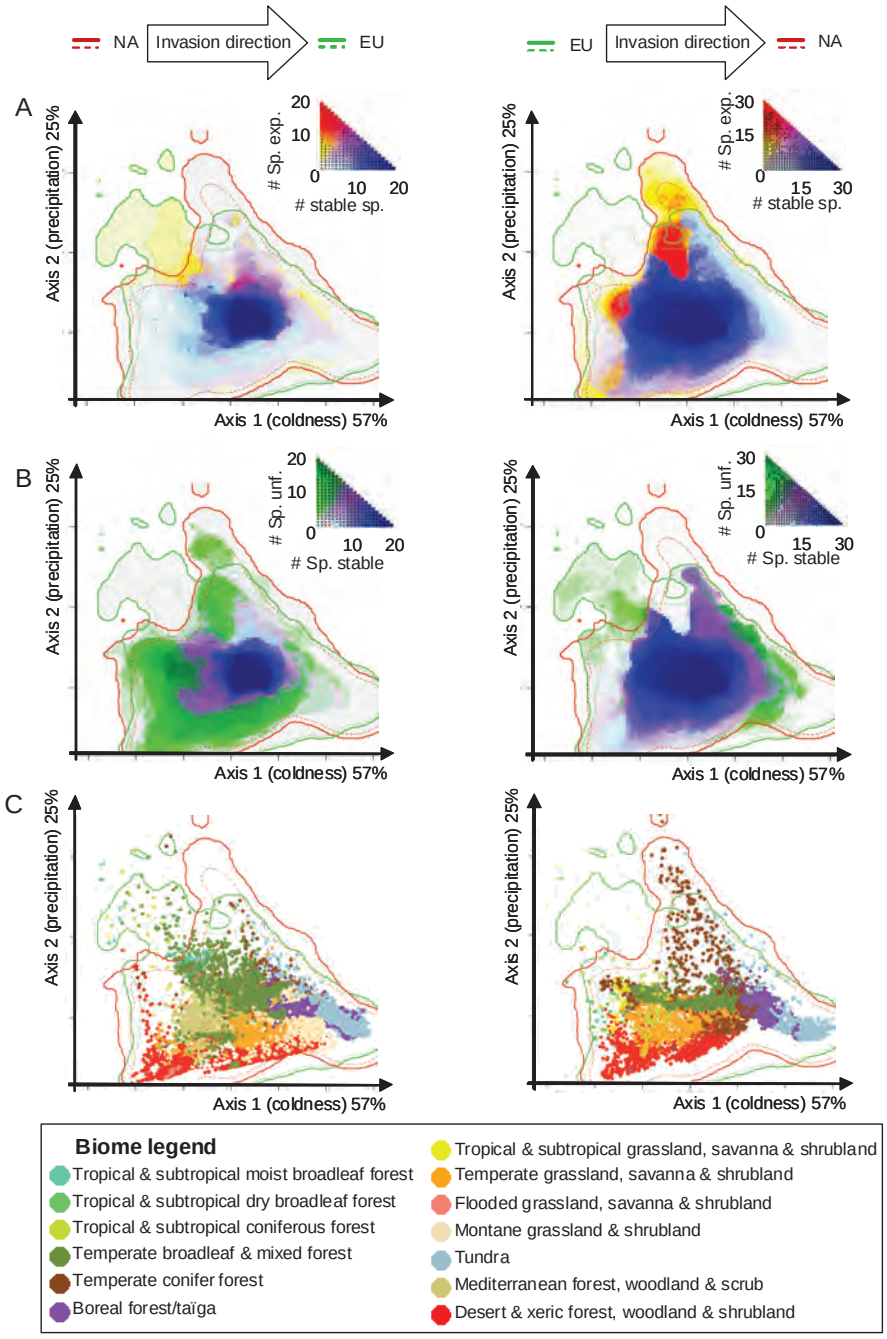


Figure 3: Niche dynamic between native and invaded ranges in Holarctic environmental space depicted by the first two axes of a principal component analysis, calibrated on the entire range of conditions available in NA (red contour lines) and EU (green contour lines). Niche expansion, overlap and unfilling situations are stacked in the environmental space for each species. Bidimensional color keys represent the number of species showing expansion (grey to red, A), unfilling (grey to green, B) and overlap (grey to blue, A and B). Occupied color classes are shown by black dots. C represents the distribution of biomes in the invaded environmental space.

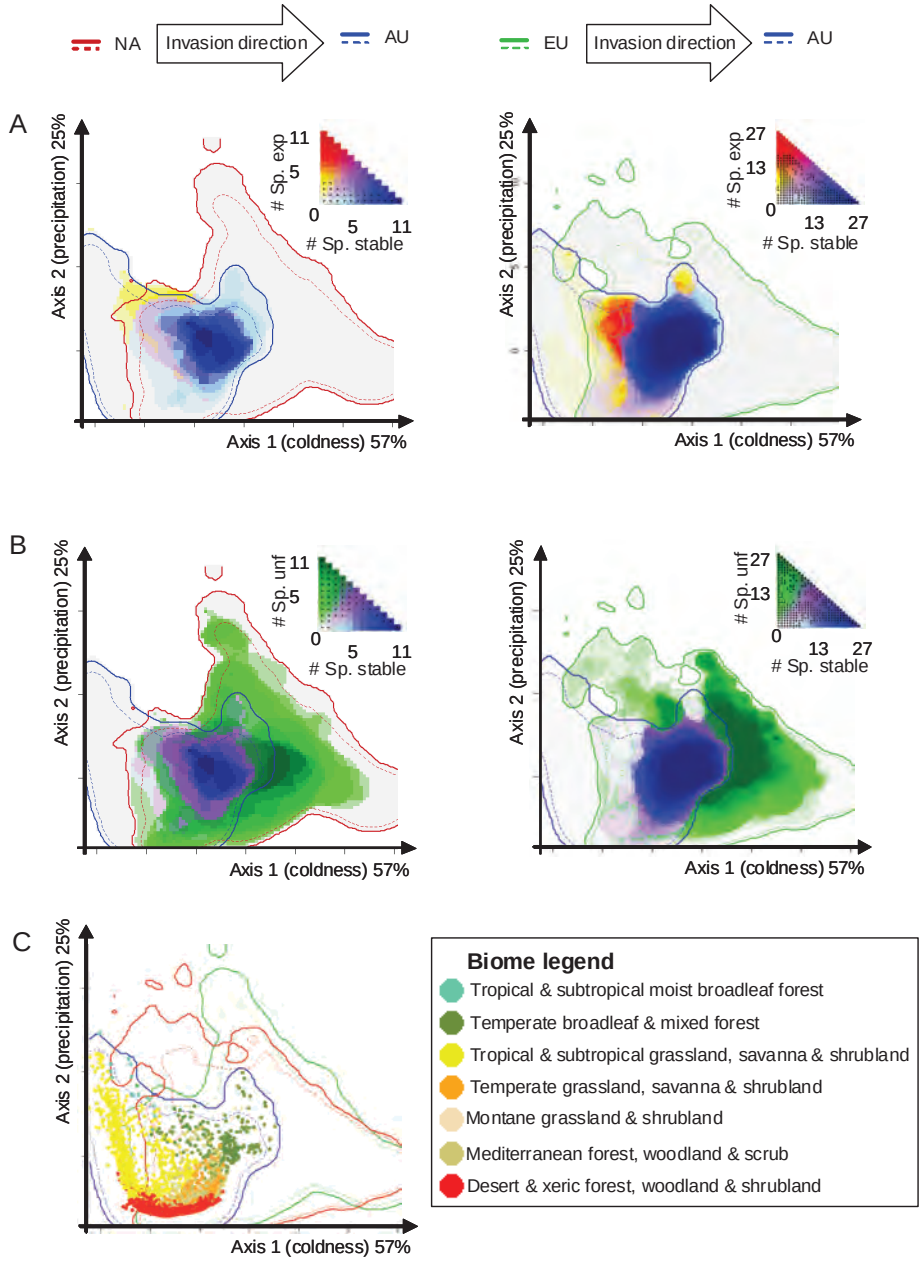


Figure 4: Niche dynamic between native and invaded ranges in Australian environmental space. Same legend as Fig. 3, except realized environment in AU is additionally represented (blue contour lines) and C represents biomes distribution in AU.

Our findings have implications for anticipating biological invasions. They suggest that ENMs remain reasonable tools to predict invasions from the native range if study areas have comparable environments, at least with regard to the biologically relevant variables. This was indeed the rule in most of our species and thus is likely to also apply to many other terrestrial alien invasive plants. To illustrate this, we built ENM for each species' native distribution. The models reveal on average a fair transferability, with only a minority of poor predictions in the invaded range (8 NA species and 2 EU species) based on the Boyce index (B; Supp. Mat.). As expected, we found a positive correlation between B and the niche overlap D, and negative correlations between B and total niche changes (Fig. S6). Interestingly, similar results are obtained when comparing niche metrics with ENM

predictions calibrated on the analog climates between EU and NA or on the whole climate (Fig. S7). Using the approach to niche comparison (Broennimann *et al.*, 2011) as a complement to ENMs thus remains important because it allows disentangling of disequilibrium situations, such as niche expansion or partial filling, in analog climates (Fig. 1).

Our findings that climatic niche shifts are rare among terrestrial plant invaders between their native and introduced ranges parallels results from a recent study showing that increase in species' abundance are rare between ranges (Firn *et al.*, 2011). We found only a few plant invaders (e.g., spotted knapweed) showing an important proportion of their invaded range outside their native niche, possibly resulting from ecological and/or evolutionary changes, although we cannot exclude dispersal limitation in the native range as a possible contributing factor. Conversely, most reported niche differences are likely caused by partial filling of the native niche in the invaded range. Recognizing that some cases of true niche change do exist, further assessments should seek to understand strategies that have allowed these particular alien invasive species to expand their niches dramatically, with possible implications for biocontrol (Muller-Scharer *et al.,* 2004). Although our study focused on Holarctic plant invaders, they included a wide range of plants, ranging from trees to herbs. It would be particularly interesting to use the same framework to test whether the same pattern is found in other organisms, especially in aquatic plants, as some of these are known to have a very large invaded range compared to their native one (Forno, 1983). Finally, our study specifically tested for niche change between geographic regions, but our general finding of niche conservation also supports an important role for ENMs in assessments of species vulnerability to climate change over time (Pearman *et al.*, 2008).

ACKNOWLEDGMENTS

We thank Robin Engler, Eva Schumacher and Courtney Angelo who contributed to the compilation of the data; Jake Alexander, Hugh Possingham, Loïc Pellissier and Sarah M. Gray for their comments on the manuscript. AG, OB and BP received main support from the National Center for Competence in Research « Plant Survival ». CD and CK received support from the USDA National Institute of Food and Agriculture, Biology of Weedy and Invasive Species Program Grant no. 2006-35320-17360. Climatic data used in this study are available Supp. Mat. Species data used in this study were assembled from multiple sources with various release politics. All data sources are described in the SOM.

Supporting Supp. Mat Material for

Climatic niche shifts are rare among terrestrial plant invaders

Blaise Petitpierre, Christoph Kueffer, Olivier Broennimann, Christophe Randin, Curtis Daehler and Antoine Guisan

correspondence to: antoine.guisan@unil.ch

This supplement includes:

References and notes

MATERIALS AND METHODS

Data

Our species selection was based on different sources and expert judgment. It represents many of the major invaders of EU and NA while also representing different life forms (herbs, trees etc.). We also included many of the most anciently introduced EU species in NA. Some of these are mainly agricultural weeds rather than natural area invaders, but we included them because they experienced a particularly long presence in NA (> 300 years). Occurrence data were obtained through the GBIF Data Portal (data.gbif.org, accessed on the 06/16/2009) for every species. However, GBIF data are often partial and biased. Species distributions were then completed throughout different sources and databases (Table S1, Table S2 and Table S3). Note that we used the taxonomy from the Integrated Taxonomic Information System (ITIS http://www.itis.gov, retrieved on the 09/30/11) except for *Aster novi-belgii* L. in place of *Symphyotrichum novi-belgii* var. *novi-belgii* (L.) Nesom. Maps from atlases were scanned and georeferenced and only occurrences with accuracy finer or equal to 50 km were kept. This resolution is coarse but commonly used in global environmental modeling studies and it represents the best compromise to: (i) cover the full range of species, which is a necessary condition to get an unbiased assessment of niche conservatism, and (ii) be able to use enough observations to get a robust analytical power. It was particularly important to sample the native niche as exhaustively as possible to cover all possible environments where the species can grow. Even with exhaustive data, a species in its native range may not grow in all suitable environments due to dispersal limitations (Svenning and Skov, 2004)

The importance of the extents (boundaries of area) chosen in ecological niche modeling has major implications in the delimitation of the niche (Barve *et al.*, 2011). For accurate comparison, it is especially important that the extent represents area that has been accessible to the species of interest over relevant time periods (Barve *et al.*, 2011). The three different extents of the study areas were delimited following Cox's new definition of the biogeographic kingdoms for plants (Cox, 2001) for North America (NA) and Eurasia (EU). Coastlines were used to delimit the extent in Australia (AU). Biogeographic kingdoms are coarse units but they include historical, climatic and phylogeographical components and they allow extents to be kept in all comparisons.

Climatic niche was depicted by eight climatic variables (Table S4, New *et al.*, 1999, *41*) at a resolution of 0.5° (approximately 50 km.). Among these variables, we derived growing degree days, potential evapotranspiration and ratio of actual and potential evapotranspiration following Thuiller *et al.* (Thuiller *et al.*, 2005). Thus each species and biome (Olson *et al.*, 2001) occurrence was related to climatic conditions. The coarse resolution of the analyses could bias results toward showing fewer niche shifts, as fine-scale shifts may not be detectable at this resolution. Yet it is the resolution most commonly used at this scale and the best that can currently be used from a practical standpoint. We

screened the literature to find the oldest known date of introduction in the Holarctic invaded range for every species when this information was available (Table S5).

General Approach

The general approach is summarized in Fig. S2. We used the same two-step framework proposed by Broennimann *et al*. (Broennimann *et al.*, 2011) to quantify niche changes between native and invaded ranges. First, the realized niche is represented in a gridded plan depicted by the first two components of a principal component analysis (PCA) calibrated using climatic conditions (Table S4) present inside the whole Holarctic study extent (i.e. NA and EU). The first axis mostly correlates with temperature and the second with precipitation variables (Fig. S3). We did not include more axes because they only explain a marginal proportion of the total variation. However, when niche overlap is measured including the third PCA axis, it is correlated to the overlap measured on the two first axes (Table S7). Among numerous ordination or ENM's techniques, this approach has been shown to be the most accurate to assess niche overlap, at least between two Holarctic range (Broennimann *et al.*, 2011). In this global environmental space, species occurrences were converted into species occurrence densities (Broennimann *et al.*, 2011). Second, this environmental space was gridded and occurrences were converted into densities. A kernel function was applied to smooth the distribution of the densities. The whole environment, i.e., all the available sites in the study area, was also converted into densities. Thus, climatic availability can be taken into account by correcting species densities with environmental densities. Broennimann *et al*. (Broennimann *et al.*, 2011) referred to these corrected densities with the term "species occupancy". When the native range was compared with AU, AU sites were projected along the axes of the PCA calibrated on EU and NA sites. Thus, the environmental space remains constant allowing a more robust comparison

Niche overlap and test of niche conservatism

We measured niche overlap between native and invaded ranges using the same approach as Broennimann *et al*. (Broennimann *et al.*, 2011). The index we used was Schoener's D, measured on the occupancies in the environmental space depicted by the two first axes of the PCA (Broennimann *et al.*, 2011, Warren *et al.*, 2008). Schoener's D indicates the overall match between two niches over the whole climatic space and varies between 0 (no overlap at all) and 1 (complete overlap). Niche occupancy corresponds to species density corrected by the environmental availability (Broennimann *et al.*, 2011). Thus, we tested whether the overlap between native and invaded distribution was higher compared to a random distribution in the invaded range, taking climatic availability into account (Broennimann *et al.*, 2011, Warren *et al.*, 2008). This is a one-sided version of the test proposed by Broennimann *et al*. (Broennimann *et al.*, 2011).

Niche unfilling and expansion

Schoener's D on species occupancy disentangles climate availability but does not take into account the difference between partial filling and expansion. When native and invaded ranges overlapped in climatic space, three categories were considered: (1) stable environments where species occurs in both ranges (Fig. S1 A), (2) unfilled environments where species occur only in the native range (Fig. S1 B), and (3) new environments where the species occur only in the invaded range (Fig. S1 C). Indices quantifying these three categories were calculated as follows. Stability is the proportion of the densities in the alien distribution overlapping with native distribution. Unfilling is the proportion of the densities in the native distribution located in different conditions than the alien distribution. Expansion is the proportion of the densities in the alien distribution located in different conditions than the native distribution (or 1-stability). Crucially, these indices are based only on the climate space that is shared between the native and invaded range. Therefore expansions measured in this way characterize true niche shifts, e.g., caused by rapid evolution or changes in biotic interactions, although dispersal limitation in the native range could also contribute. These niche change indices were calculated at the intersection of the dashed line (75% percentile) in Fig. 3 and Fig. 4. We used a percentile approach to remove the marginal environments to avoid biases due to some side effects of the kernel smoothing. However, sensitivity analyses revealed that using different percentiles (75, 80, 85, 90, 95 and 100%) to delimit the marginal environments (100% = including all marginal climates) had no effect on the niche change metrics, except for unfilling in Australia (Kruskal-Wallis rank sum test, p-val < 0.001) (Fig. S5) due to the fact that most of the Australian climate does not overlap with the Holarctic conditions. In the main analyses, we therefore used the 75^{th} percentile of site density in each extent as a threshold to remove marginal environments. For comparison, we also measured niche change indices based on the total climates (Fig. S7).

Methodology matters to quantify niche shifts

Methodology matters to disentangle the nature of different niche changes and measure the extent of niche shifts (Broennimann *et al.*, 2011, Peterson, 2011, Rodder and Lotters, 2009, Warren *et al.*, 2008, Wiens and Graham, 2005*4*). In contrast to our study, expansion was previously reported for 18 out of 26 (69%) species in Australia (Gallagher *et al.*, 2010). However, this higher frequency of niche expansion mainly results from methodological differences. Following Broennimann *et al*. (Broennimann *et al.*, 2007) and Gallagher *et al*. (Gallagher *et al.*, 2010) we quantified the difference of the niche position in the environmental space with a principal component analysis calibrated on environmental conditions at each species occurrence. Using the R CRAN library 'ade4', we calculated a between-class inertia ratio for each native/invasive comparison to assess the magnitude and the significance of the difference between native and invasive distributions. Species distributions were weighted such that native and invasive data had the same weight in the analysis. As in Gallagher *et al*. (Gallagher *et al.*, 2010) we qualitatively assessed the

overlap of the convex hulls encompassing occurrence in the environmental space depicted by the PCA bi-plots into the five categories: near-complete overlap, alien range is a subset, native range is a subset, partial overlap and complete dislocation. Then we qualified the three categories: native range is a subset, partial overlap and complete dislocation as expansion (Fig. S9 and S10, Table S8). When applying this earlier approach to our dataset, we also found a significant difference in niche position (inertia-ratio of a between-class PCA analysis) for every comparison in EU, NA and AU, with 38 (76%) species showing apparent expansion in the Holarctic and 20 (53%) species showing apparent expansion in AU (Table S8, Fig. S6 and S7). We consider these apparent shifts (or pseudo-shifts) as ecologically non-meaningful for two reasons. Firstly, because it a priori maximizes the inter-group variance, this statistical approach has recently been shown to overestimate niche shifts (Broennimann *et al.*, 2011). Secondly, differences in climatic availability between the two ranges need to be accounted for to correctly quantify species niche shifts (Broennimann *et al.*, 2011, Mandle *et al.*, 2010, Peterson, 2011*3*). In our study, we used novel technical developments addressing both issues and find little evidence of significant niche expansion associated with invasion of new regions.

Ecological niche modeling

ENMs use environmental conditions at locations where species occur to depict the species' realized niche (Guisan and Thuiller, 2005), i.e. the envelope of conditions where the species can thrive after accounting for dispersal limitations and biotic interactions (Soberon, 20075). ENM projections therefore strongly rely on the realized niche being conserved between native and invaded ranges (Pearman *et al.*, 2008). To fit ENM, we used the same species distributions and climate data and the same environmental extent than in the niche analysis, i.e. we removed the non-analogous climates detected through the PCA to avoid extrapolation of models. As different statistical techniques can produce different results (Elith *et al.*, 2006*6*), we used an ensemble of four different techniques: artificial neural networks (ANN, Ripley, 1996*7*), boosted regression trees (BRT, Friedman *et al.*, 2000), generalized linear models (GLM) and maximum entropy modeling (MAXENT, Phillips *et al.*, 2006). We used the BIOMOD library (Thuiller *et al.*, 2009), available for the R software, for the modeling with ANN, BRT and GLM with all the default parameters. MAXENT was used with the dismo library for R, also with default parameters. Every 50 km cell that was not a presence was considered as a pseudo-absence to calibrate statistical models requiring presences and absences and as background data in MAXENT. Models were calibrated with 70% of the data in the native range (training dataset) and then projected onto the remaining 30% of the native range (evaluation dataset), the reciprocal Holarctic invaded range (EU or NA) and the Australian range (AU). For each technique, presences and pseudo-absences used to calibrate the model were weighted such as to ensure neutral (0.5) prevalence. The procedure was replicated 10 times, with random training and evaluation datasets, such that we obtained 40 models (10 replicates x 4 techniques) per species.

We used the Boyce index (Boyce *et al.*, 2002, Hirzel *et al.*, 2006) to assess model performance. Contrary to common evaluation measures for presence-absence data, like AUC (Fielding and Bell, 1997) or TSS (Allouche *et al.*, 2006) which present problems for presence-only data or for comparisons between species when the extents of calibration area differ (see Lobo *et al.*, 2008, Lobo and Tognelli, Phillips *et al.*, 2006), the Boyce index only requires presences and measures how much model predictions differ from random distribution of the observed presences across the prediction gradients (Boyce *et al.*, 2002). It is thus the most appropriate metric in the case of presence-only models. It is continuous and varies between -1 and +1. Positive values indicate a model which present predictions are consistent with the distribution of presences in the evaluation dataset, values close to zero mean that the model is not different from a random model, negative values indicate counter predictions, i.e., predicting poor quality areas where presences are more frequent (Hirzel *et al.*, 2006). We calculated the Boyce index for each of the 40 models per species and averaged the values to obtain a final evaluation for the 30% left-out data in the native range (B_{nat}), in the invaded range in Holarctic (i.e. EU or NA, B_{inv}) and in AU when data were available (B_{au}).

Overall, EU models show better transferability than NA models (Table S9, Fig. S6 and S7; average B_{inv} for NA invaders models = 0.10 +/-0.26; EU invaders = 0.45 +/-0.23, t-test for difference, $p < 0.001$; average B_{au} for NA invaders models in Australia = 0.43 +/-0.26; EU invaders = 0.60 +/-0.16, t-test for difference, $p = 0.06$). This difference can be explained by (i) our partial filling results, as it is a far more important component of the apparent niche differences than niche expansion and is especially affecting NA species, or (ii) the choice of predictors (Mandle *et al.*, 2010, Peterson, 2011). If expansion is detected, one alternative solution is to fit ENMs on pooled species dataset, merging occurrences in the native and invaded ranges, so as to fit the broadest possible realized niche for the species (Beaumont *et al.*, 2009, Broennimann and Guisan, 2008).

Boyce index significantly correlates with niche unfilling and total niche changes (i.e. unfilling + expansion, correlation test based on Spearman's ρ, Fig. S6) in Holarctic and Australia. However, expansion does not seem to significantly affect models predictions (Fig. S6), likely because they a rare events. Schoener's D also correlates with the quality of the models predictions in the invaded range and is a good index of models transferability. Interestingly, there is no drastic difference when models are calibrated on the whole geographical extent, i.e. including the non-analogous climates (Fig. S7). However, when the climatic conditions are more different, (e.g. between Holarctic and Australia), the trends are not significant (Fig S7 E-H).

Authors contributions

AG and CK initiated the idea. OB and CR further helped designing the initial study. AG and CD supported the data collection, and CK, BP and OB coordinated it. OB and BP wrote all analytical codes and, with CK, improved the design further. BP ran all analyses and

prepared all figures with major inputs from OB, CK and AG, and to a lesser extent from CR and CD. All authors helped interpreting the results. BP wrote the first paper draft with major help from OB, CK and AG. All authors contributed to writing the final manuscript.

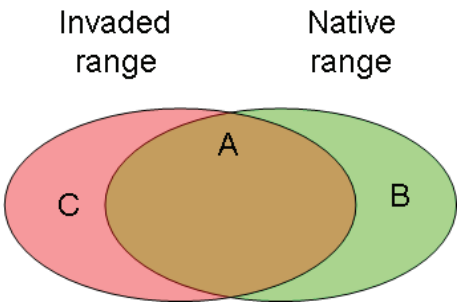


Fig. S1: Graphical representation in environmental space of the three categories obtained by a niche comparison between two ranges: stable (A, the part of the niche where the species occurs in both ranges), unfilled (B, where the species occurs in the native range only) and expansion (C where the species newly occurs in the invaded range). These three categories are used to calculate the stability, unfilling and expansion indices, respectively (see main text and supplementary methods for details).

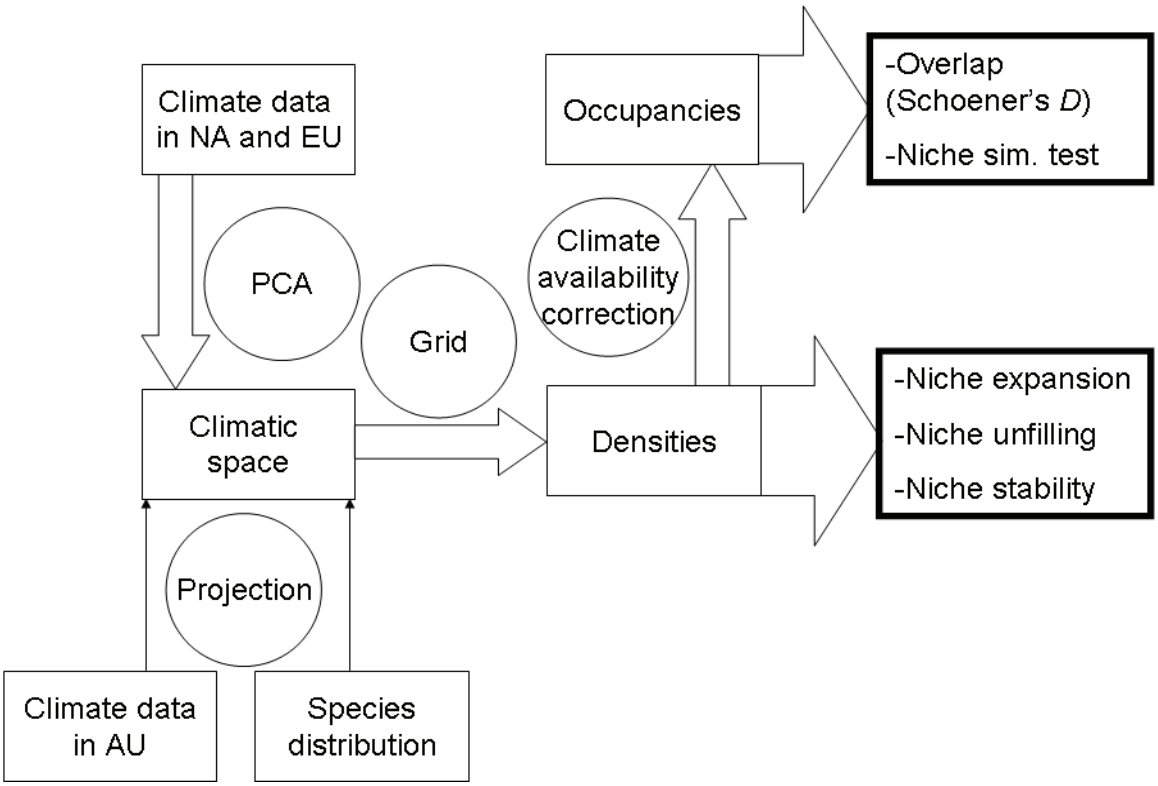


Fig. S2: Scheme of the general approach.

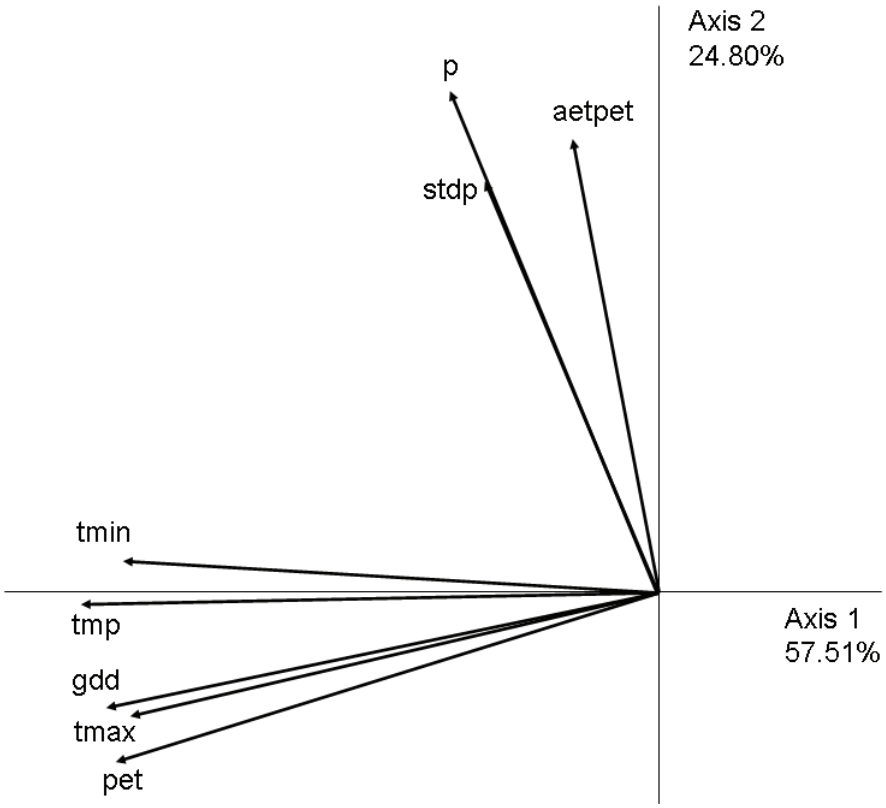
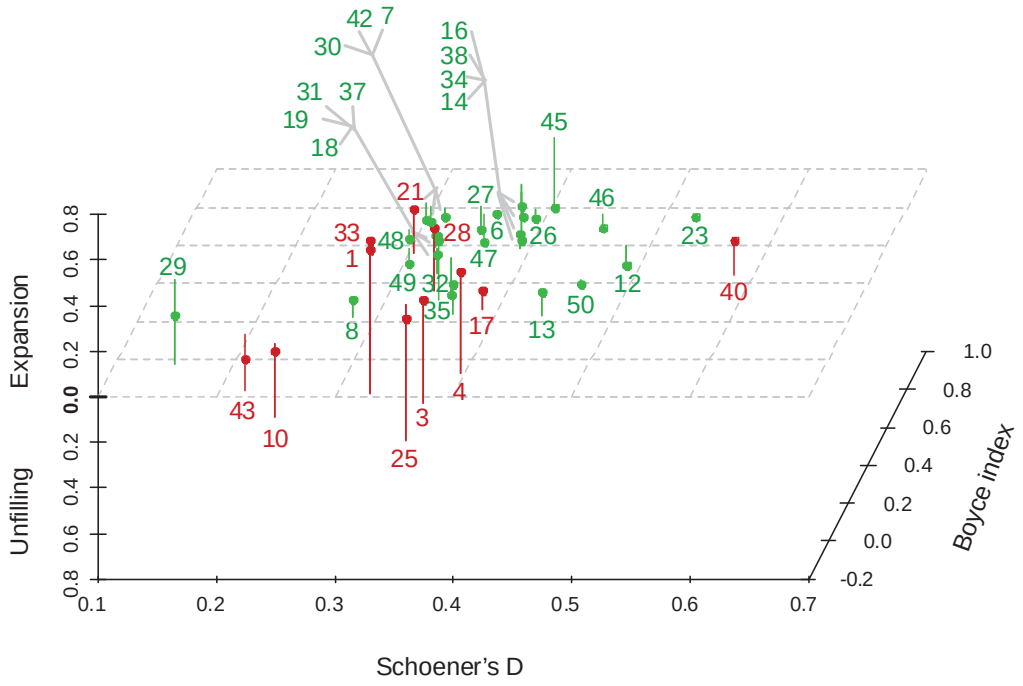


Fig. S3: Correlation between climatic variables and the first two components of the principal component analysis calibrated on the climatic conditions in North America (NA) and Eurasia (EU). First and second components explain 82.31% of the total climatic variation in NA and EU. The abbreviations of variables are described in Table S4.



1	*Acer negundo*	19	*Dactylis glomerata*	35	*Plantago major*
3	*Amaranthus retroflexus*	21	*Epilobium ciliatum*	36	*Poa annua*
4	*Ambrosia artemisiifolia*	23	*Erodium cicutarium*	37	*Potentilla recta*
6	*Anagallis arvensis*	25	*Helianthus tuberosus*	40	*Robinia pseudoacacia*
7	*Anthoxanthum odoratum*	26	*Holcus lanatus*	42	*Rumex acetosella*
8	*Arabidopsis thaliana*	27	*Hypochaeris radicata*	43	*Solidago canadensis*
10	*Aster novi-belgii*	28	*Juncus tenuis*	45	*Sonchus oleraceus*
12	*Bromus sterilis*	29	*Linaria vulgaris*	46	*Trifolium arvense*
13	*Bromus tectorum*	30	*Lythrum salicaria*	47	*Trifolium dubium*
14	*Carduus nutans*	31	*Medicago lupulina*	48	*Trifolium repens*
16	*Cirsium vulgare*	32	*Melilotus albus*	49	*Verbascum thapsus*
17	*Conyza canadensis*	33	*Phytolacca americana*	50	*Vicia sativa*
18	*Cytisus scoparius*	34	*Plantago lanceolata*		

Fig. S4: Niche changes between Holarctic native and invaded ranges in AU. Segments represent the magnitude of niche changes for each species. Extension above and below the zero plane indicate expansion and unfilling, respectively. Intersections with the zero plane are shown with dots. Green and red colors indicate Eurasian and North American species origins respectively. Niche change indices are plotted against niche overlap (i.e., Schoener's D) and the Boyce index (i.e., the evaluation of ENM's calibrated in the native range and projected onto analog climates in the invaded range).

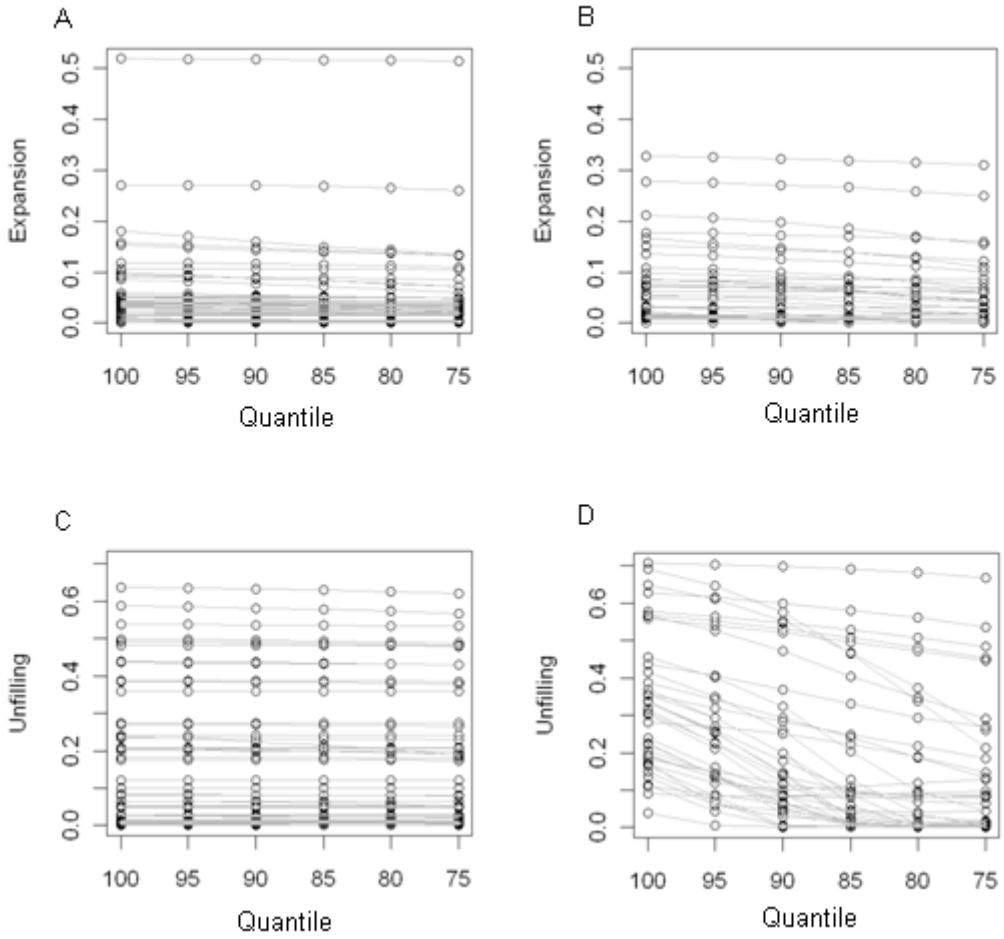


Fig. S5: Sensitivity of expansion (A,B) and unfilling (C,D) indices in Holarctic (A,C) and Australia (B,D) to the inclusion of marginal climates. Six different quantiles of densities were used as a threshold to remove marginal environments in the principal component analysis. Only unfilling in Australia (D) is significantly affected by the choice of the threshold to remove marginal climates (Kruskall-Wallis rank sum test, p-val < 0.001).

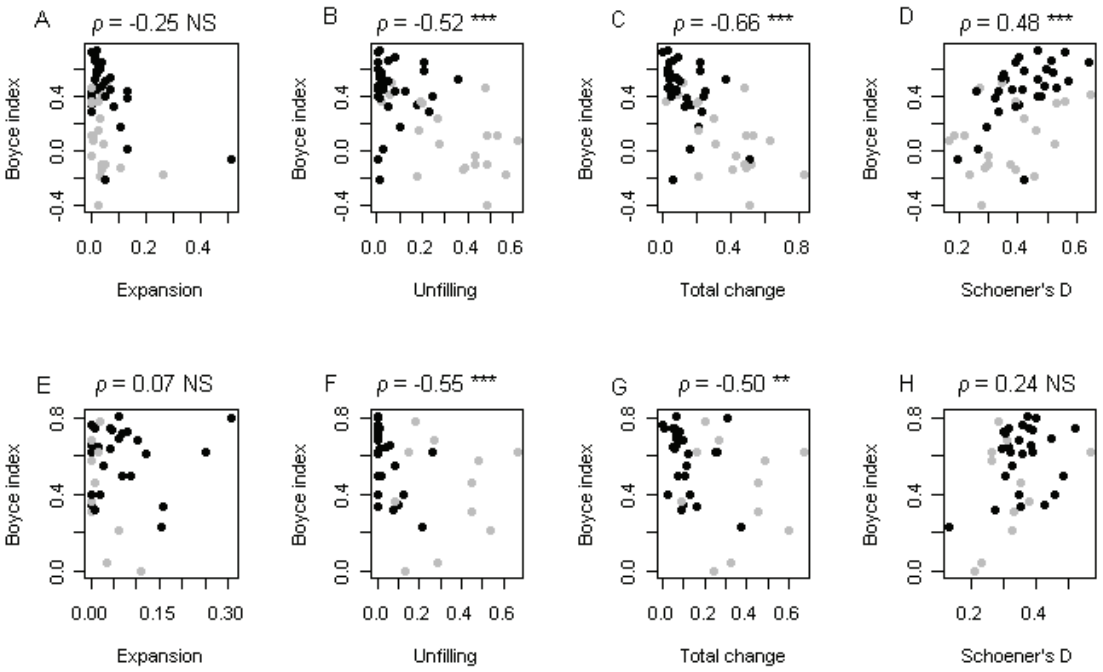


Fig. S6: Correlation between models performance in the invaded ranges (Holarctic, A-D; Australian E-H) and niche changes. Correlation between the Boyce index, expansion, unfilling, total change (i.e. expansion + unfilling) and Schoener's D was assessed with correlation test based on Spearman's ρ. Black and grey dots represent species origin (EU and NA respectively).

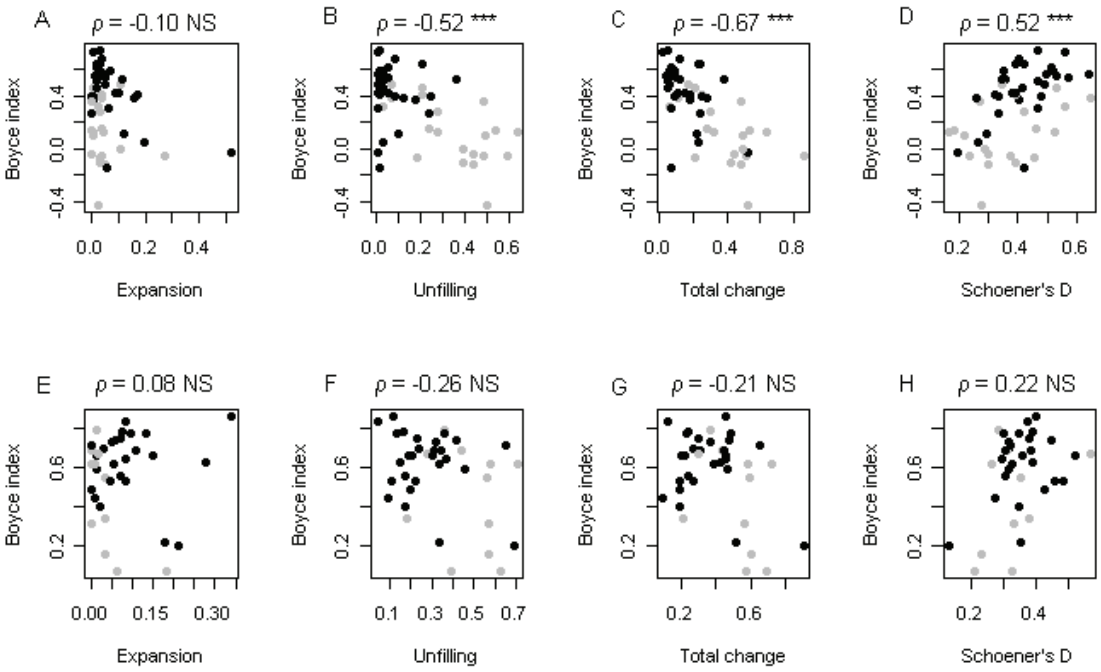


Fig. S7: Same legend as Fig. S6 except that models were calibrated on the entire native geographical range, as the niche changes were measured on the whole environmental extent (without removing marginal climates).

Fig S8 is downloadable at:

https://docs.google.com/open?id=0B5ZZgCyoOEdnMjA1OWY4NTctYWRlZi00NmFmLTgzMDUtZjRjMmJlODVjMzkz

Fig. S8: Environmental niche of native and alien species distributions. Environmental space is depicted by the two first components of a principal component analysis calibrated on the climatic conditions in North America and Eurasia. The first line represents the distribution of extent (grey dots) and occurrences (colored dots) for native (green) alien in Holarctic (red) and alien in Australia (blue) distributions. The second line shows these distributions converted into densities and the three categories of species environments in native, alien in Holarctic and alien in Australia distributions: unfilling (green), stability (blue) and expansion (red). Dashed lines delimit marginal climates that were not taken into account in the analyses.

Fig S9 is downloadable at:

https://docs.google.com/open?id=0B5ZZgCyoOEdnZjIzYTU1ODYtZmY0YS00OGYxLWE4NTgtMDg4MmY3ZDRjMzYy

Fig. S9: Comparison of the native (green) and Holarctic alien (red) niche based on a principal component analysis calibrated on species occurrences. The minimum convex polygons that encompassed all the occurrence points for each species were used to classify species into the five groups presented in Table S7. The correlation circle shows how variables are correlated to the two first axes. A between-class analysis tested the magnitude and significance of the shift between occurrence clouds, yielding a between-class inertia ratio, which was tested for significance using 999 Monte Carlo randomizations.

Fig S10 is downloadable at:

https://docs.google.com/open?id=0B5ZZgCyoOEdnY2MyYmQ3N2UtNzdhNi00MDFlLWFmNjUtOWI0MWZjYjZiOTA2

Fig. S10: Comparison of the native (green) and Australian alien (blue) niche based on a principal component analysis calibrated on species occurrences. The minimum convex polygons that encompassed all the occurrence points for each species were used to classify species into the five groups presented in Table S7. The correlation circle show how variables are correlated to the two first axes. A between-class analysis tested the magnitude and significance of the shift between occurrence clouds, yielding a between-class inertia ratio, which was tested for significance using 999 Monte Carlo randomizations.

Table S1: Complete species name and sources used for Eurasian distributions are map atlases (Hultén and Fries, 1986, Meusel and Jaeger, 1992) and Supp. Mat database: http://www.agroatlas.ru (Afonin et al., 2008). Use of the source is indicated by the symbol "+".

Species	Hulten-Fries	Meusel-Jaeger	agroatlas.ru	others
Acer negundo L.	-	-	-	-
Alliaria petiolata (M. Bieb.) Cavara & Grande	+	-	-	-
Amaranthus retroflexus L.	+	-	+	-
Ambrosia artemisiifolia L.	-	+	+	-
Amorpha fruticosa L.	-	-	-	-
Anagallis arvensis L.	+	-	-	-
Anthoxanthum odoratum L.	+	-	-	-
Arabidopsis thaliana (L.) Heynh.	-	-	-	(Hoffmann, 2002)
Asclepias syriaca L.	-	-	+	-
Aster novi-belgii L.	-	-	-	-
Bidens frondosa L.	-	+	-	-
Bromus sterilis L.	+	-	-	-
Bromus tectorum L.	+	-	-	-
Carduus nutans L.	-	+	+	-
Centaurea stoebe L	-	+	-	Personal data
Cirsium vulgare (Savi) Ten.	+	-	-	-
Conyza canadensis (L.) Cronquist	-	+	+	-
Cytisus scoparius (L.) Link	+	-	-	-
Dactylis glomerata L.	+	-	-	-
Echinocystis lobata (Michx.) Torr. & Gray	-	+	-	-
Epilobium ciliatum Raf.	+	-	-	-
Erigeron annuus (L.) Pers.	-	+	-	-
Erodium cicutarium (L.) L'Hér. ex Ait.	+	-	+	-
Euphorbia esula L.	+	-	-	-
Helianthus tuberosus L.	-	+	-	-
Holcus lanatus L.	+	-	-	-
Hypochaeris radicata L.	+	-	-	-
Juncus tenuis Willd.	+	-	-	-
Linaria vulgaris P. Mill.	+	-	-	-
Lythrum salicaria L.	+	-	-	-
Medicago lupulina L.	+	-	-	-
Melilotus albus Medik.	+	-	-	-
Phytolacca americana L.	-	-	-	-
Plantago lanceolata L.	+	-	+	-
Plantago major L.	+	-	+	-
Poa annua L.	+	-	+	-
Potentilla recta L.	+	-	-	-
Prunus serotina Ehrh.	-	-	-	Personal data
Rhus hirta (L.) Sudworth	-	-	-	-
Robinia pseudoacacia L.	-	-	-	Personal data
Rudbeckia laciniata L.	-	+	-	-
Rumex acetosella L.	+	-	+	-
Solidago canadensis L.	-	+	-	-
Solidago gigantea Aiton	-	+	-	-
Sonchus oleraceus L.	+	-	-	-
Trifolium arvense L.	+	-	-	-
Trifolium dubium Sibth.	+	-	-	-
Trifolium repens L.	+	-	+	-
Verbascum thapsus L.	+	-	-	-
Vicia sativa L.	+	-	-	-

Table S2: Species name and sources used for North American distributions are map atlases (Hultén and Fries, 1986, Meusel and Jaeger, 1992, Little, 1971) and Supp. Mat database: IPANE (Mehrhoff *et al.*, 2003), Consortium of California Herbaria (CCH, ucjeps.berkeley.edu/consortium/ consulted on the 06/18/2009), Vegbank (http://vegbank.org/ consulted on the 05/12/2009) and AKEPIC (http://aknhp.uaa.alaska.edu/maps/akepic/, consulted on the 05/18/2009). Use of the source is indicated by the symbol "+".

Species	Hulten-Fries	Meusel-Jaeger	Little	IPANE	CCH	Vegbank	AKEPIC	other
Acer negundo	-	-	+	-	-	-	-	-
Alliaria petiolata	-	-	-	+	+	+	+	(Welk *et al.*, 2002)
Amaranthus retroflexus	+	-	-	-	-	+	-	-
Ambrosia artemisiifolia	-	+	-	-	+	-	+	Personal data
Amorpha fruticosa	-	-	-	+	+	+	-	-
Anagallis arvensis	+	-	-	-	-	-	-	-
Anthoxanthum odoratum	+	-	-	-	+	+	+	-
Arabidopsis thaliana	-	-	-	-	-	-	-	(Hoffmann, 2002)
Asclepias syriaca	-	-	-	-	-	-	-	-
Aster novi-belgii	-	+	-	-	-	-	-	-
Bidens frondosa	-	+	-	-	-	-	-	-
Bromus sterilis	+	-	-	+	+	+	-	-
Bromus tectorum	+	-	-	+	+	+	+	-
Carduus nutans	-	-	-	-	+	+	-	-
Centaurea stoebe	-	-	-	+	+	+	+	Personal data
Cirsium vulgare	+	-	-	-	+	+	+	-
Conyza canadensis	-	+	-	-	+	+	-	-
Cytisus scoparius	+	-	-	-	-	+	+	-
Dactylis glomerata	+	-	-	-	+	+	+	-
Echinocystis lobata	-	+	-	-	-	+	-	-
Epilobium ciliatum	+	-	-	-	-	-	-	-
Erigeron annuus	-	+	-	-	+	+	-	-
Erodium cicutarium	+	-	-	-	+	+	+	-
Euphorbia esula	+	-	-	+	+	+	-	-
Helianthus tuberosus	-	+	-	-	-	-	-	-
Holcus lanatus	+	-	-	-	+	+	+	-
Hypochaeris radicata	+	-	-	-	+	+	+	-
Juncus tenuis	+	-	-	-	+	+	-	-
Linaria vulgaris	+	-	-	-	+	+	+	-
Lythrum salicaria	+	-	-	+	+	+	+	-
Medicago lupulina	+	-	-	-	+	+	+	-
Melilotus albus	+	-	-	-	+	+	+	-
Phytolacca americana	-	-	-	-	+	+	-	Personal data
Plantago lanceolata	+	-	-	-	+	+	+	-
Plantago major	+	-	-	-	+	+	+	-
Poa annua	+	-	-	-	+	+	+	-
Potentilla recta	+	-	-	-	+	+	-	-
Prunus serotina	-	-	+	-	+	+	-	Personal data
Rhus hirta	-	-	+	-	-	+	-	-
Robinia pseudoacacia	-	-	+	-	+	+	-	Personal data
Rudbeckia laciniata	-	+	-	-	-	-	-	-
Rumex acetosella	+	-	-	+	+	+	+	-
Solidago canadensis	-	+	-	-	+	+	-	Personal data
Solidago gigantea	-	+	-	-	-	+	-	Personal data
Sonchus oleraceus	+	-	-	-	+	+	+	-
Trifolium arvense	+	-	-	-	+	+	-	-
Trifolium dubium	+	-	-	-	+	+	-	-
Trifolium repens	+	-	-	-	+	+	+	-

Verbascum thapsus	+		-	-	+	+	-	-
Vicia sativa	+	-	-	-	+	+	-	-

Table S3: Species name and sources used for Australian distributions Supp. Mat databases : Australia's Virtual Herbarium (http://www.ersa.edu.au/avh/ accessed on the 01/27/2009), Department of Sustainability and Environment (DSE) of Victoria and data provided by the *Royal Botanic Gardens Board, Melbourne, MELISR database accessed on the 02/10/2009*. Use of the source is indicated by the symbol "+".

Species	**AVH**	**DSE**	**MELISR**
Acer negundo	-	-	-
Amaranthus retroflexus	+	+	+
Ambrosia artemisiifolia	+	-	+
Anagallis arvensis	+	+	+
Anthoxanthum odoratum	+	+	+
Arabidopsis thaliana	+	-	-
Aster novi-belgii	-	-	-
Bromus sterilis	+	-	+
Bromus tectorum	+	+	+
Carduus nutans	+	+	+
Cirsium vulgare	+	+	+
Conyza canadensis	+	+	+
Cytisus scoparius	+	+	+
Dactylis glomerata	+	+	+
Epilobium ciliatum	+	-	-
Erodium cicutarium	+	+	+
Helianthus tuberosus	-	-	-
Holcus lanatus	+	+	+
Hypochaeris radicata	+	+	+
Juncus tenuis	+	+	+
Linaria vulgaris	+	+	+
Lythrum salicaria	-	+	+
Medicago lupulina	+	+	+
Melilotus albus	+	+	+
Phytolacca americana	+	-	+
Plantago lanceolata	+	+	+
Plantago major	+	+	+
Poa annua	+	+	+
Potentilla recta	+	+	+
Robinia pseudoacacia	+	+	+
Rumex acetosella	+	+	+
Solidago canadensis	+	-	+
Sonchus oleraceus	+	+	+
Trifolium arvense	+	+	+
Trifolium dubium	+	+	+
Trifolium repens	+	+	+
Verbascum thapsus	+	+	+
Vicia sativa	+	+	+

Table S4: List of climatic variables.

Variable	Description
aetpet	Ratio of actual to potential evapotranspiration
pet	Potential evapotranspiration
p	Annual amount of precipitations
std_p	Annual variation of precipitations
tmin	Minimum temperature of the coldest month
tmp	Annual mean temperature
tmax	Maximum temperature of the warmest month
gdd	Growing degree-days above 5 °C

Table S5: Niche change indices and minimum residence time between North America (NA) and Eurasia (EU). The native range (Nat. range), the overlap between native and invaded range (Schoener's D), the p-value of the niche similarity test (Sim. test), niche expansion (E), stability (S), unfilling (U) and minimum residence time (MRT), i.e. elapsed time in years between 1st known introduction and 2009 are given for each species.

Species	Nat. range	D	Sim. test		E	S	U	MRT
Acer negundo	NA	0.278	0.157	NS	0.025	0.975	0.489	310
Alliaria petiolata	EU	0.404	0.003	**	0.002	0.998	0.174	141
Amaranthus retroflexus	NA	0.645	0.050	*	0.041	0.959	0.053	259
Ambrosia artemisiifolia	NA	0.392	0.007	**	0.024	0.976	0.189	246
Amorpha fruticosa	NA	0.235	0.107	NS	0.260	0.740	0.567	285
Anagallis arvensis	EU	0.335	0.001	***	0.002	0.998	0.079	359
Anthoxanthum odoratum	EU	0.356	0.023	*	0.016	0.984	0.359	209
Arabidopsis thaliana	EU	0.351	0.022	*	0.022	0.978	0.013	no data
Asclepias syriaca	NA	0.372	0.022	*	0.037	0.963	0.482	359
Aster novi-belgii	NA	0.456	0.002	**	0.031	0.969	0.181	230
Bidens frondosa	NA	0.531	0.043	*	0.001	0.999	0.196	273
Bromus sterilis	EU	0.294	0.082	NS	0.109	0.891	0.099	219
Bromus tectorum	EU	0.533	0.025	*	0.015	0.985	0.023	129
Carduus nutans	EU	0.479	0.132	NS	0.004	0.996	0.241	149
Centaurea stoebe	EU	0.194	0.132	NS	0.513	0.487	0.000	116
Cirsium vulgare	EU	0.468	0.111	NS	0.014	0.986	0.027	289
Conyza canadensis	NA	0.563	0.001	***	0.003	0.997	0.021	363
Cytisus scoparius	EU	0.263	0.124	NS	0.133	0.867	0.030	159
Dactylis glomerata	EU	0.518	0.187	NS	0.019	0.981	0.023	no data
Echinocystis lobata	NA	0.390	0.130	NS	0.036	0.964	0.379	103
Epilobium ciliatum	NA	0.185	0.062	NS	0.000	1.000	0.534	118
Erigeron annuus	NA	0.343	0.005	**	0.059	0.941	0.064	374
Erodium cicutarium	EU	0.465	0.054	NS	0.022	0.978	0.014	312
Euphorbia esula	EU	0.572	0.035	*	0.050	0.950	0.050	149
Helianthus tuberosus	NA	0.290	0.143	NS	0.105	0.895	0.386	382
Holcus lanatus	EU	0.324	0.094	NS	0.133	0.867	0.011	309
Hypochaeris radicata	EU	0.344	0.125	NS	0.072	0.928	0.008	130
Juncus tenuis	NA	0.219	0.020	*	0.006	0.994	0.482	185
Linaria vulgaris	EU	0.415	0.173	NS	0.019	0.981	0.029	327
Lythrum salicaria	EU	0.337	0.019	*	0.004	0.996	0.230	195
Medicago lupulina	EU	0.516	0.138	NS	0.014	0.986	0.048	219
Melilotus albus	EU	0.643	0.060	NS	0.036	0.964	0.006	345
Phytolacca americana	NA	0.463	0.032	*	0.028	0.972	0.187	309
Plantago lanceolata	EU	0.493	0.236	NS	0.041	0.959	0.006	279
Plantago major	EU	0.562	0.034	*	0.002	0.998	0.004	337
Poa annua	EU	0.496	0.286	NS	0.030	0.970	0.002	212
Potentilla recta	EU	0.395	0.106	NS	0.018	0.982	0.206	109

Prunus serotina	NA	0.167	0.102	NS	0.005	0.995	0.621	386
Rhus hirta	NA	0.523	0.001	***	0.043	0.957	0.273	407
Robinia pseudoacacia	NA	0.421	0.069	NS	0.034	0.966	0.264	375
Rudbeckia laciniata	NA	0.300	0.012	*	0.050	0.950	0.430	359
Rumex acetosella	EU	0.420	0.276	NS	0.049	0.951	0.011	379
Solidago canadensis	NA	0.272	0.004	**	0.000	1.000	0.480	364
Solidago gigantea	NA	0.297	0.006	**	0.000	1.000	0.430	251
Sonchus oleraceus	EU	0.381	0.250	NS	0.072	0.928	0.006	138
Trifolium arvense	EU	0.422	0.167	NS	0.020	0.980	0.208	no data
Trifolium dubium	EU	0.261	0.016	*	0.132	0.868	0.121	no data
Trifolium repens	EU	0.470	0.329	NS	0.048	0.952	0.001	359
Verbascum thapsus	EU	0.390	0.258	NS	0.085	0.915	0.052	259
Vicia sativa	EU	0.406	0.015	*	0.016	0.984	0.081	129

Table S6: Niche change indices between native range in North America (NA) or Eurasia (EU) and the invaded range in Australia. The native range (Nat. range), the overlap between native and invaded range (Schoener's D), the p-value of the niche similarity test (Sim. test), niche expansion (E), stability (S) and unfilling (U) are given for each species.

Species	**Nat. range**	**D**	**Sim. test**		**E**	**S**	**U**
Acer negundo	NA	0.264	0.003	**	0.000	1.000	0.739
Amaranthus retroflexus	NA	0.333	0.028	*	0.000	1.000	0.592
Ambrosia artemisiifolia	NA	0.351	0.008	**	0.008	0.972	0.454
Anagallis arvensis	EU	0.357	0.010	*	0.002	0.987	0.006
Anthoxanthum odoratum	EU	0.314	0.001	***	0.043	0.843	0.105
Arabidopsis thaliana	EU	0.272	0.001	***	0.009	0.984	0.394
Aster novi-belgii	NA	0.229	0.006	**	0.033	0.951	0.426
Bromus sterilis	EU	0.489	0.001	***	0.086	0.801	0.097
Bromus tectorum	EU	0.429	0.001	***	0.000	1.000	0.430
Carduus nutans	EU	0.385	0.001	***	0.000	1.000	0.380
Cirsium vulgare	EU	0.389	0.014	*	0.250	0.509	0.000
Conyza canadensis	NA	0.378	0.011	*	0.000	1.000	0.199
Cytisus scoparius	EU	0.315	0.001	***	0.009	0.964	0.168
Dactylis glomerata	EU	0.314	0.001	***	0.016	0.920	0.250
Epilobium ciliatum	NA	0.285	0.001	***	0.020	0.951	0.467
Erodium cicutarium	EU	0.526	0.002	**	0.007	0.948	0.039
Helianthus tuberosus	NA	0.326	0.017	*	0.061	0.902	0.528
Holcus lanatus	EU	0.391	0.005	**	0.046	0.846	0.089
Hypochaeris radicata	EU	0.350	0.008	**	0.101	0.687	0.000
Juncus tenuis	NA	0.309	0.001	***	0.000	1.000	0.630
Linaria vulgaris	EU	0.129	0.015	*	0.155	0.748	0.190
Lythrum salicaria	EU	0.299	0.002	**	0.078	0.719	0.019
Medicago lupulina	EU	0.324	0.004	**	0.029	0.922	0.210
Melilotus albus	EU	0.353	0.002	**	0.160	0.742	0.078
Phytolacca americana	NA	0.262	0.001	***	0.000	1.000	0.770
Plantago lanceolata	EU	0.380	0.025	*	0.043	0.831	0.051
Plantago major	EU	0.350	0.007	**	0.000	1.000	0.322
Poa annua	EU	0.374	0.001	***	0.062	0.753	0.008
Potentilla recta	EU	0.319	0.005	**	0.000	1.000	0.602
Robinia pseudoacacia	NA	0.569	0.011	*	0.016	0.943	0.231
Rumex acetosella	EU	0.304	0.008	**	0.069	0.772	0.008
Solidago canadensis	NA	0.208	0.043	*	0.110	0.819	0.216
Sonchus oleraceus	EU	0.402	0.004	**	0.309	0.620	0.000
Trifolium arvense	EU	0.451	0.011	*	0.060	0.803	0.023
Trifolium dubium	EU	0.358	0.003	**	0.122	0.654	0.000

Trifolium repens	EU	0.292	0.011	*	0.042	0.871	0.155
Verbascum thapsus	EU	0.304	0.002	**	0.069	0.750	0.020
Vicia sativa	EU	0.457	0.001	***	0.018	0.921	0.100

Table S7: Pearson correlation coefficients between Schoener's D when measured on the first two axes (D_{12}), on the second and third axes (D_{23}) and on the first and third axes (D_{13}) of the principal component analysis.

	D_{12}	D_{23}
D_{23}	0.58	
D_{13}	0.75	0.68

Table S8: Between-class PCA inertia ratio (IR) and categories of overlap between native and invaded range in Holarctic (HO) and Australia (AU).

Species	Nat. range	IR HO	IR AU	Categorie nat-HO	Categorie nat-AU
Acer negundo	NA	0.23	0.16	partial overlap	inv. is a subset
Alliaria petiolata	EU	0.09	NA	partial overlap	no data
Amaranthus retroflexus	NA	0.13	0.22	partial overlap	inv. is a subset
Ambrosia artemisiifolia	NA	0.11	0.37	partial overlap	inv. is a subset
Amorpha fruticosa	NA	0.23	NA	inv. is a subset	no data
Anagallis arvensis	EU	0.09	0.29	inv. is a subset	inv. is a subset
Anthoxanthum odoratum	EU	0.18	0.40	inv. is a subset	partial overlap
Arabidopsis thaliana	EU	0.15	0.26	partial overlap	inv. is a subset
Asclepias syriaca	NA	0.36	NA	partial overlap	no data
Aster novi-belgii	NA	0.17	0.27	partial overlap	partial overlap
Bidens frondosa	NA	0.12	NA	inv. is a subset	no data
Bromus sterilis	EU	0.17	0.27	partial overlap	partial overlap
Bromus tectorum	EU	0.07	0.18	partial overlap	inv. is a subset
Carduus nutans	EU	0.11	0.29	inv. is a subset	inv. is a subset
Centaurea stoebe	EU	0.14	NA	nat. is a subset	no data
Cirsium vulgare	EU	0.15	0.42	partial overlap	partial overlap
Conyza canadensis	NA	0.05	0.22	partial overlap	partial overlap
Cytisus scoparius	EU	0.12	0.25	partial overlap	inv. is a subset
Dactylis glomerata	EU	0.09	0.33	partial overlap	partial overlap
Echinocystis lobata	NA	0.14	NA	partial overlap	no data
Epilobium ciliatum	NA	0.19	0.20	inv. is a subset	inv. is a subset
Erigeron annuus	NA	0.14	NA	partial overlap	no data
Erodium cicutarium	EU	0.16	0.34	partial overlap	inv. is a subset
Euphorbia esula	EU	0.12	NA	inv. is a subset	no data
Helianthus tuberosus	NA	0.26	0.33	partial overlap	inv. is a subset
Holcus lanatus	EU	0.12	0.29	partial overlap	partial overlap
Hypochaeris radicata	EU	0.14	0.35	partial overlap	partial overlap
Juncus tenuis	NA	0.19	0.20	partial overlap	inv. is a subset
Linaria vulgaris	EU	0.20	0.48	partial overlap	partial overlap
Lythrum salicaria	EU	0.16	0.36	inv. is a subset	partial overlap
Medicago lupulina	EU	0.12	0.35	partial overlap	partial overlap
Melilotus albus	EU	0.09	0.41	partial overlap	partial overlap
Phytolacca americana	NA	0.23	0.30	partial overlap	inv. is a subset
Plantago lanceolata	EU	0.10	0.35	partial overlap	partial overlap
Plantago major	EU	0.04	0.35	partial overlap	inv. is a subset
Poa annua	EU	0.11	0.39	partial overlap	partial overlap
Potentilla recta	EU	0.25	0.28	partial overlap	inv. is a subset
Prunus serotina	NA	0.32	NA	partial overlap	no data

Rhus hirta	NA	0.32	NA	partial overlap	no data
Robinia pseudoacacia	NA	0.27	0.22	partial overlap	inv. is a subset
Rudbeckia laciniata	NA	0.24	NA	partial overlap	no data
Rumex acetosella	EU	0.24	0.42	partial overlap	partial overlap
Solidago canadensis	NA	0.14	0.42	inv. is a subset	partial overlap
Solidago gigantea	NA	0.17	NA	inv. is a subset	no data
Sonchus oleraceus	EU	0.12	0.45	partial overlap	partial overlap
Trifolium arvense	EU	0.22	0.40	partial overlap	partial overlap
Trifolium dubium	EU	0.25	0.37	partial overlap	partial overlap
Trifolium repens	EU	0.15	0.42	partial overlap	inv. is a subset
Verbascum thapsus	EU	0.20	0.38	partial overlap	partial overlap
Vicia sativa	EU	0.18	0.33	inv. is a subset	inv. is a subset

Table S9: Synthesis of the evaluations of ENM. Values of the Boyce index (B) are given for the native range (B_{nat}, where models are calibrated), the invaded range in Holarctic (B_{inv}) and, if data were available, in the invaded range in Australia (B_{au}). 'No data' indicate that no data were available for the species in Australia and thus the evaluation metrics could not be computed.

Species	**B_{nat}**	**B_{inv}**	**B_{au}**
Acer negundo	0.945	-0.403	0.579
Alliaria petiolata	0.884	0.341	no data
Amaranthus retroflexus	0.935	0.412	0.31
Ambrosia artemisiifolia	0.94	0.364	0.457
Amorpha fruticosa	0.626	-0.179	no data
Anagallis arvensis	0.853	0.434	0.761
Anthoxanthum odoratum	0.908	0.528	0.745
Arabidopsis thaliana	0.861	0.564	0.314
Asclepias syriaca	0.84	-0.097	no data
Aster novi-belgii	0.751	-0.182	0.04
Bidens frondosa	0.882	0.356	no data
Bromus sterilis	0.882	0.173	0.494
Bromus tectorum	0.862	0.469	0.349
Carduus nutans	0.957	0.406	0.661
Centaurea stoebe	0.851	-0.066	no data
Cirsium vulgare	0.823	0.531	0.624
Conyza canadensis	0.894	0.363	0.364
Cytisus scoparius	0.872	0.02	0.65
Dactylis glomerata	0.815	0.581	0.648
Echinocystis lobata	0.946	-0.132	no data
Epilobium ciliatum	0.944	0.112	0.785
Erigeron annuus	0.854	0.507	no data
Erodium cicutarium	0.954	0.74	0.747
Euphorbia esula	0.924	0.511	no data
Helianthus tuberosus	0.811	-0.126	0.215
Holcus lanatus	0.874	0.396	0.737
Hypochaeris radicata	0.916	0.541	0.682
Juncus tenuis	0.946	0.117	0.687
Linaria vulgaris	0.831	0.448	0.231
Lythrum salicaria	0.895	0.293	0.731
Medicago lupulina	0.94	0.672	0.551
Melilotus albus	0.959	0.648	0.338
Phytolacca americana	0.779	0.153	0.622
Plantago lanceolata	0.918	0.48	0.744

Plantago major	0.916	0.731	0.396
Poa annua	0.953	0.604	0.808
Potentilla recta	0.889	0.655	0.626
Prunus serotina	0.85	0.078	no data
Rhus hirta	0.82	0.054	no data
Robinia pseudoacacia	0.785	0.236	0.621
Rudbeckia laciniata	0.889	-0.098	no data
Rumex acetosella	0.917	-0.208	0.72
Solidago canadensis	0.96	0.467	-0.005
Solidago gigantea	0.792	-0.035	no data
Sonchus oleraceus	0.879	0.452	0.8
Trifolium arvense	0.909	0.588	0.694
Trifolium dubium	0.846	0.446	0.616
Trifolium repens	0.965	0.406	0.635
Verbascum thapsus	0.82	0.327	0.499
Vicia sativa	0.927	0.692	0.397

REFERENCES

Afonin, A.N., Greene, S.L., Dzyubenko, N.I. & Frolov, A.N. (2008) http://www.agroatlas.ru.

Alexander, J.M. & Edwards, P.J. (2010) Limits to the niche and range margins of alien species. *Oikos,* **119**, 1377-1386.

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

Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling,* **222**, 1810-1819.

Beaumont, L.J., Gallagher, R.V., Thuiller, W., Downey, P.O., Leishman, M.R. & Hughes, L. (2009) Different climatic envelopes among invasive populations may lead to underestimations of current and future biological invasions. *Diversity and Distributions,* **15**, 409-420.

Blair, A.C. & Hufbauer, R.A. (2009) Hybridization and invasion: one of North America's most devastating invasive plants shows evidence for a history of interspecific hybridization. *Evolutionary Applications,* **3**, 40-51.

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

Broennimann, O., Fitzpatrick, M.C., Pearman, P., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N., Graham, C. & Guisan, A. (2011) Changes in ecological niches in space and time: biological signal or artifact? *Global Ecology and Biogeography*.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters,* **4**, 585-589.

Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters,* **10**, 701-709.

Callaway, R.M., Waller, L.P., Diaconu, A., Pal, R., Collins, A.R., Mueller-Schaerer, H. & Maron, J.L. Escape from competition: Neighbors reduce *Centaurea stoebe* performance at home but not away. *Ecology,* **92**, 2208-2213.

Cox, B.C. (2001) The biogeographic regions reconsidered. *Journal of Biogeography,* **28**, 511-523.

Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P. & Linder, H.P. (2009) Phylogenetic biome conservatism on a global scale. *Nature,* 458, 754-U90.

Crosby, A.W. (1986) *Ecological Imperialism: The Biological Expansion of Europe, 900-1900 (Studies in Environment and History).* Cambridge University Press.

Durka, W., Bossdorf, O., Prati, D. & Auge, H. (2005) Molecular evidence for multiple introductions of garlic mustard (Alliaria petiolata, Brassicaceae) to North America. *Molecular Ecology,* **14**, 1697-1706.

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

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.

Firn, J., Moore, J.L., MacDougall, A.S., Borer, E.T., Seabloom, E.W., HilleRisLambers, J., Harpole, W.S., Cleland, E.E., Brown, C.S., Knops, J.M.H., Prober, S.M., Pyke, D.A., Farrell, K.A., Bakker, J.D., O'Halloran, L.R., Adler, P.B., Collins, S.L., D'Antonio, C.M., Crawley, M.J., Wolkovich, E.M., La Pierre, K.J., Melbourne, B.A., Hautier, Y., Morgan, J.W., Leakey, A.D.B., Kay, A., McCulley, R., Davies, K.F., Stevens, C.J., Chu, C.-J., Holl, K.D., Klein, J.A., Fay, P.A., Hagenah, N., Kirkman, K.P. & Buckley, Y.M. (2011) Abundance of introduced species at home predicts abundance away in herbaceous communities. *Ecology Letters,* **14**, 274-281.

Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography,* **16**, 24-33.

Forno, I.W. (1983) Native distribution of the Salvinia auriculata complex and keys to species identification. *Aquatic Botany,* **17**, 71-83.

Friedman, J., Hastie, T. & Tibshirani, R. (2000) Additive logistic regression: A statistical view of boosting. *Annals of Statistics,* **28**, 337-374.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology,* **98**, 790-799.

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

Hierro, J.L., Villarreal, D., Eren, O., Graham, J.M. & Callaway, R.M. (2006) Disturbance facilitates invasion: The effects are stronger abroad than at home. *American Naturalist,* **168**, 144-156.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling,* **199**, 142-152.

Hoffmann, M.H. (2002) Biogeography of *Arabidopsis thaliana* (L.) Heynh. (*Brassicaceae*). *Journal of Biogeography,* **29**, 125-134.

Hultén, E. & Fries, M. (1986) *Atlas of north European vascular plants north of the tropic of Cancer.* Koeltz Scientific Books, Königstein.

Klironomos, J.N. (2002) Feedback with soil biota contributes to plant rarity and invasiveness in communities. *Nature,* **417**, 67-70.

Lauzeral, C., Leprieur, F., Beauchard, O., Duron, Q., Oberdorff, T. & Brosse, S. (2011) Identifying climatic niche shifts using coarse-grained occurrence data: a test with non-native freshwater fish. *Global Ecology and Biogeography,* **20**, 407-414.

Lavergne, S. & Molofsky, J. (2007) Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America,* **104**, 3883-3888.

Leung, B., Lodge, D.M., Finnoff, D., Shogren, J.F., Lewis, M.A. & Lamberti, G. (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **269**, 2407-2413.

Little, E.L.J. (1971) *Atlas of United States trees, volume 1, conifers and important hardwoods: U.S. Department of Agriculture Miscellaneous Publication.*

Lobo, J.M., Jimenez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography,* **17**, 145-151.

Lobo, J.M. & Tognelli, M.F. Exploring the effects of quantity and location of pseudo-absences and sampling biases on the performance of distribution models with limited point occurrence data. *Journal for Nature Conservation,* **19**, 1-7.

Mandle, L., Warren, D.L., Hoffmann, M.H., Peterson, A.T., Schmitt, J. & von Wettberg, E.J. (2010) Conclusions about niche expansion in introduced *Impatiens walleriana* populations depend on method of analysis. *Plos One,* **5**, 9.

Medley, K.A. (2010) Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus* Skuse (*Culicidae*), revealed by reciprocal distribution models. *Global Ecology and Biogeography,* **19**, 122-133.

Mehrhoff, L.J., Silander, J.A., Jr., Leicht, S.A., Mosher, E.S. & Tabak, N.M. (2003) http://www.ipane.org.

Meusel, H. & Jaeger, E. (1992) *Vergleichende Chorologie der zentraleuropäischen Flora.* Fischer, Jena.

Muller-Scharer, H., Schaffner, U. & Steinger, T. (2004) Evolution in invasive plants: implications for biological control. *Trends in Ecology & Evolution,* **19**, 417-422.

New, M., Hulme, M. & Jones, P. (1999) Representing twentieth-century space-time climate variability. Part I: Development of a 1961-90 mean monthly terrestrial climatology. *Journal of Climate,* 12, 829-856.

Novak, S.J. & Mack, R.N. (2001) Tracing Plant Introduction and Spread: Genetic Evidence from Bromus tectorum (Cheatgrass). *Bioscience,* **51**, 114-122.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. (2001) Terrestrial Ecoregions of the World: A New Map of Life on Earth. *Bioscience,* **51**, 933-938.

Pairon, M., Petitpierre, B., Campbell, M., Guisan, A., Broennimann, O., Baret, P.V., Jacquemart, A.-L. & Besnard, G. (2010) Multiple introductions boosted genetic diversity in the invasive range of black cherry (Prunus serotina; Rosaceae). *Annals of Botany,* **105**, 881-890.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008) Niche dynamics in space and time. *Trends in Ecology & Evolution,* **23**, 149-158.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography,* **38**, 817-827.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* **190**, 231-259.

Rejmanek, M. (2000) Invasive plants: approaches and predictions. *Austral Ecology,* **25**, 497-506.

Ripley, B.D. (1996) *Pattern Recognition and Neural Networks.* Cambridge University Press.

Rodder, D. & Lotters, S. (2009) Niche shift versus niche conservatism? Climatic characteristics of the native and invasive ranges of the Mediterranean house gecko (*Hemidactylus turcicus*). *Global Ecology and Biogeography,* **18**, 674-687.

Seastedt, T.R. & Pyšek, P. (2011) Mechanisms of plant invasions of North America and European grasslands. *Annual Review of Ecology, Evolution, and Systematics,* 42.

Seipel, T., Kueffer, C., Rew, L.J., Daehler, C.C., Pauchard, A., Naylor, B.J., Alexander, J.M., Edwards, P.J., Parks, C.G., Arevalo, J.R., Cavieres, L.A., Dietz, H., Jakobs, G., McDougall, K., Otto, R. & Walsh, N. (2011) Processes at multiple scales affect richness and similarity of non-native plant species in mountains around the world. *Global Ecology and Biogeography*, no-no.

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

Svenning, J.C. & Skov, F. (2004) Limited filling of the potential range in European tree species. *Ecology Letters,* **7**, 565-573.

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

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (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.

Treier, U.A., Broennimann, O., Normand, S., Guisan, A., Schaffner, U., Steinger, T. & Mueller-Schaerer, H. (2009) Shift in cytotype frequency and niche space in the invasive plant *Centaurea maculosa*. *Ecology,* **90**, 1366-1377.

Vilà, M., Basnou, C., Pyšek, P., Josefsson, M., Genovesi, P., Gollasch, S., Nentwig, W., Olenin, S., Roques, A., Roy, D. & Hulme, P.E. (2009) How well do we understand the impacts of alien species on ecosystem services? A pan-European, cross-taxa assessment. *Frontiers in Ecology and the Environment,* **8**, 135-144.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. *Evolution,* **62**, 2868-2883.

Welk, E. (2004) Constraints in range predictions of invasive plant species due to non-equilibrium distribution patterns: Purple loosestrife (*Lythrum salicaria*) in North America. *Ecological Modelling,* **179**, 551-567.

Welk, E., Schubert, K. & Hoffmann, M.H. (2002) Present and potential distribution of invasive garlic mustard (*Alliaria petiolata*) in North America. *Diversity and Distributions,* **8**, 219-233.

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

Woodward, F.I. (1987) *Climate and plant distribution.* Cambridge University Press, Cambridge, UK.

Xu, C.-Y., Julien, M.H., Fatemi, M., Girod, C., Van Klinken, R.D., Gross, C.L. & Novak, S.J. (2009) Phenotypic divergence during the invasion of Phyla canescens in Australia and France: evidence for selection-driven evolution. *Ecology Letters,* **13**, 32-44.

CHAPTER 1.3

The importance of variable selection in species distribution models: the case of biological invasions

Blaise Petitpierre[1], Olivier Broennimann[1], Christoph Kueffer [2,3], Curtis Daehler[2] & Antoine Guisan[1]

[1] Department of Ecology & Evolution, University of Lausanne, 1015 Lausanne, Switzerland

[2] Department of Botany, University of Hawaii at Manoa, 3190 Maile Way, Honolulu, HI 96822, USA

[3] Institute of Integrative Biology, ETH Zürich, Universitätstrasse 16, 8092 Zürich, Switzerland

This paper is in preparation for *Methods in Ecology and Evolution*

My contribution to the project: I gathered the data, initiated the question, ran the analysis and wrote the initial draft of the paper.

ABSTRACT

Niche-based models of species distribution (SDM) are widely used to predict impacts of human induced global change on biodiversity, in particular for anticipating biological invasions. However, reliable predictions rely on SDM transferability, which itself strongly depends on the choice of predictors in the models. Here, we use a systematic approach to evaluate, based on presences only, which predictor selection approach provides the most robust SDM predictions.

We use the invaded ranges of 50 Holarctic plant invaders, previously shown to exhibit niche stability between ranges, as independent datasets to evaluate SDM calibrated in their native range. We tested 10 different strategies to select predictors from an initial set of 27 macroclimatic predictors, accounting variable for issues of predictors' proximality, multicollinearity and analogy.

We found SDMs to be transferable for 96% of species, with best combination of predictors being species-specific. Still, two variables strategies proved better SDM transferability overall: choosing a set of predictors known for their proximal effects on plant distribution, and using only the two first axes of a principal component analysis calibrated on all climatic predictors. We also found that internal cross-validation was not sufficient to inform about SDM transferability.

We conclude that anticipating the invasion potential of the main Holarctic plant invaders is possible at the macroclimatic scale, but strongly rely on the choice of predictors to be included in the models. More generally, our study proposed an approach to assess variable selection in SDMs and usefully inform about the best strategies to use for transferring SDMs between ranges, with potential applications to predict models over time.

INTRODUCTION

Species distribution models (SDM) quantify estimates of ecological niches by relating observed species occurrences to environmental variables. Projections of SDM onto the geographical space thus draw potential distributions of species (Guisan and Zimmermann, 2000). SDMs are routinely used in many fields of ecology, evolution and conservation (Guisan and Thuiller, 2005). One of the most popular applications of SDM is the projection of potential species distribution in different geographic areas or time periods (i.e. an issue of SDM transferability). Projections in space may be used to represent the potential distribution in other geographical areas that a species reached naturally (e.g. different mountain chains; Randin *et al.*, 2006, Pellissier *et al.*, 2013a) or with the help of human activities (e.g. invasive species; Peterson, 2003, Thuiller *et al.*, 2005). SDM can also be projected back in time (hindcasting, e.g. to depict potential glacial refugia; Pearman *et al.*, 2008*a*, Maiorano *et al.*, 2012) or into the future (forecasting, e.g. to assess the impact of climate change on biodiversity; Engler *et al.*, 2011) following climate change scenarios. These approaches are especially useful for supporting biological conservation in an era

where biodiversity is massively put at threat by human activities (Jeschke and Strayer, 2008, Guisan *et al*. 2013).

However, some SDMs fail to predict species distribution onto different areas (e.g. Randin *et al.*, 2006, Broennimann *et al.*, 2007, Zanini *et al.*, 2009) or onto different climatic conditions (e.g. Pearman *et al.*, 2008*a*). One major assumption when projecting SDMs is that the modeled niche remains constant through time and/or space (Peterson, 2003). This brings us to the classical theoretical differentiation between the "realized" and the "fundamental" niche (Hutchinson, 1957). The fundamental niche depicts the physiological limits of species survival whereas the realized niche depicts conditions where the species can survive after accounting for negative biotic interactions (competitors, diseases; Hutchinson's original criteria) and (in a more recent definition) limiting dispersal ability (Soberon, 2007). This niche conservatism assumption may be violated at both the level of the realized niche (e.g. changes in biotic interactions) and the fundamental niche (e.g. evolution resulting in new physiological abilities, Pearman *et al.*, 2008*b*), thus preventing robust distribution predictions by SDMs. As the realized niche is directly defined along the environmental variables used in SDM, the realized niche stasis (allowing SDM transferability) or shifts (preventing SDM transferability) strongly relies on the choice of the variables establishing the environmental space. Note that disentangling the effects of fundamental niche shifts on SDM predictions are difficult as there is no way to easily predict how evolution can affect species environmental tolerance. One has to consider at least four critical aspects when building a set of predictor variables to project potential species distribution in time or space: i) proximality, ii) multicollinearity, iii) climatic availability and analogy, and (iv) model complexity and overfitting.

Using *proximal* variables, which can define species' physiological limits, is expected to bring the modeling closer to the real requirements of the species, thus allowing more robust predictions (Austin, 2002, Kearney and Porter, 2009, Rodder *et al.*, 2009, Williams *et al.*, 2012). However, without *a priori* knowledge about species ecology and physiology, picking the most proximal variables is not obvious as they may be confounded with other, highly correlated variables. For example, it is easy to understand that temperature is more proximal than elevation, but it is more difficult to know if extreme temperatures at the cold limits (e.g. frost for some species) matter more than mean temperatures (e.g. growing-degree days) to control species' ranges (Lenz *et al.*, In Press).

Multicollinearity, when two or more variables are correlated, can significantly decrease SDM predictions accuracy on independent datasets (Graham, 2003, Dormann *et al.*, 2008, Braunisch *et al.*, 2013). A common rule of thumb is to avoid correlations between variables where the Pearson's correlation $|r|$ is higher than a fixed threshold (Green, 1979, generally 0.7 or beyond). When several variables are correlated, one should choose the most proximal to species' ecology (Guisan and Zimmermann, 2000, Austin, 2007 Austin and Van Niel, 2011, Williams *et al.*, 2012). The number of variables included in a model can be a problem as each variable represents an additional parameter in the model. Over parameterization can lead to modeling spurious interactions between environmental and

biological variable (depending on the model algorithm) without any ecological and causal relationship, thus potentially reducing transferability (Warren and Seifert, 2011). A common solution is the empirical rule of “1 in 10” (Harrell *et al.*, 1984), i.e. the use of maximum one predictor for ten species occurrences.

Next, one has to take into account the distribution of environmental variables across the whole study area(s) and more particularly whether they are similarly *available* and *analog* (i.e. same range of values) between areas and time periods (environmental anisotropy, Soberón and Peterson, 2011). For example, Greenhouse frog colonized colder temperatures in its invaded ranges that do not exist in its native range (Rodder and Lotters, 2010), thus native calibrated models should be extrapolated in the invaded range only with caution. It has been argued that SDM are only reliable under shared analog environment, i.e. in analogous environments (Fitzpatrick and Hargrove, 2009, Mandle *et al.*, 2010). Additionally, interactions between environmental factors can be different between two study area(s), e.g. two distinct ecotypes of plant species can grow in two different precipitation ranges (arctic and alpine) but share a similar temperature range (Pellissier *et al.*, 2013a). The non-analog variables (in the example, precipitation) can be removed so that the model is calibrated only on the analogous variables.

Finally, one has also to care for possible overfitting, if fitting a model with too many predictors for the number of observations available to fit the model (Randin *et al.*, 2006).

Fully testing the ability of SDMs to predict species distribution through space or time requires an independent test dataset (Guisan and Zimmermann, 2000, Araujo and Guisan, 2006). The usual split-sample approach, repeatedly and randomly leaving out a certain proportion of data within the study area to evaluate models accuracy (i.e. internal cross-validation), can be insufficient in this regard (Araujo and Guisan, 2006, Phillips *et al.*, 2006). Independent data sets are thus optimal when they are geographically or temporally separated from the training data set (Randin *et al.*, 2006, Araújo and Rahbek, 2006). Systems with a temporal separation include ancient distribution data set such as pollen fossil data (e.g. Martinez-Meyer and Peterson, 2006, Pearman *et al.*, 2008*a*, Maiorano *et al.*, 2012). Geographical separation can be achieved using two distinct study areas within the natural/native species distribution (e.g. Randin *et al.*, 2006, Pellissier *et al.*, 2013b) or native and introduced ranges of invasive species (e.g. Petitpierre *et al.*, 2012).

Yet, even though one can test model transferability through time by hindcasting, it remains formally impossible to validate projections under future scenarios. However, much on climate change impacts can be learned from biological invasions (Caplat *et al.*, 2013), as these are often connected components of global changes affecting species distributions and biodiversity (Hellmann *et al.*, 2008, Pyke *et al.*, 2008, Mainka and Howard, 2010). In this way, invasive species can be seen as an analogous test of transferability compared to rapid climate change studies since they happen in abruptly changed settings, with a high chance of altered biotic interactions. In particular, the invaded range represents a unique opportunity to test which set of variables should be included in SDMs to obtain robust

projections through space, allowing a possible analogy for SDM transferability through time in the context of rapid climate change.

When aiming at transnational pre-border weed risk assessment (see e.g. Venette *et al.*, 2010), climatic niche matching between native and potentially invaded ranges is commonly done at a coarse macroclimatic resolution (e.g. 0.5° in Thuiller *et al.*, 2005, Koop *et al.*, 2012). At the continental scale, climate is often considered the dominant driver of species distribution (Woodward, 1987) and macroclimatic studies allows generalizations about primary climatic relationships, not clouded by local complexities. In addition to climate matching between native and invaded range, species traits is another possibly important factor to explain invasions success (Van Kleunen *et al.*, 2010 but see Thompson and Davis, 2011). Coupling species traits with analysis of SDM transferability could allow a wider understanding of of invasion success and improve predictions and management success (Moles *et al.*, 2008, Gallagher *et al.*, 2010, Hulme and Barrett, 2013).

In this study, we use native and invaded ranges of the 50 Holarctic plant species used in Petitpierre *et al.* (2012) to investigate the impact of variables selection on SDM transferability. Results obtained are not only important to gain a better understanding of invasive species distributions, but may also inform on the validity of transferring SDMs across the future in a rapid and abrupt climate change context, as both phenomena are interconnected. More specifically, we ask the following questions:

- Are there optimal strategies to select variables to optimize model transferability in regard to variable proximality, collinearity and analogy?

- Which variables are the most important to optimize transferability?

- What can be the effect of variables selection on projected distributions under future climates?

- Are there any traits linked to SDM transferability?

As SDM is currently the most widely used tool to assess various global change threats to biodiversity (Guisan *et al.*, 2013), assessing their transferability is a crucial task.

METHODS

Data

We used the same distribution data as Petitpierre *et al.* (2012). It consists in the distributions of 50 Holarctic plant invaders, either native in Eurasia (EU) and invading North America (NA) or vice versa. A subset of 38 of these species was introduced in Australia (AU), which was used here as second independent invaded range outside the Holarctic (Table 1). Based on the conclusions of Petitpierre *et al.* (2012), we separated species shifting their niche (i.e. showing more than 10 % of niche expansion in analog climate) from species with stable niche. Only seven species showed niche shifts within the Holarctic ranges comparison (*A. fruticosa*, *B. sterilis*, *C. stoebe*, *C. scoparius*, *H. lanatus*,

H. tuberosus and *T. dubium*), and seven species in the Holarctic-Australian ranges comparison (*C. vulgare*, *H. radicata*, *L. vulgaris*, *M. albus*, *S. canadensis*, *S. oleraceus* and *T. dubium*). The aim of this distinction was twofold. First, models of species shifting their niche are expected to show lower performance when projected in the invaded range, whatever the variables selection used. For these species, interpreting SDM transferability should be done in the light of their particular niche dynamics between ranges. However, one can still test whether particular combinations of variables can reduce the shift and allow more successfully predictions in the invaded range to be drawn for these species.

Climatic data were downloaded from Climond database (Kriticos *et al.*, 2011). It consists of 35 bioclimatic variables at a resolution of 10 arcminutes. These variables can be split into 4 groups: temperature, precipitation, solar radiation and soil moisture indices. We didn't include the solar radiation variables as they were included in the calculation of the moisture index, as the latter should be more proximal variable for plant growth at this coarse continental scale. In total, we kept 27 variables (Table 2). We used the A1 climate change scenario for 2050 based on the CSIRO MK3.0 GCM (also available in Climond), to illustrate the impact of variables selection on projected responses to climate change. Using the library raster in the R software (version 2.15.1, R, 2005), we aggregated these data at the same resolution as the species distribution data, i.e. 0.5 degree.

Variables selection strategies

For each species, we defined 10 strategies to select variables in the SDMs (Fig. 1), implementing one or combinations of the following characteristics of the variable sets: reducing multicollinearity, reducing overfitting and increasing proximality and climate analogy in the invaded range:

1) All variables - V_{all}: We built an exhaustive model including all 27 variables available (V_{all}). This approach can be considered as a "non-strategy" to deal with any of the four characteristics of the variable sets. It has been used to predict species invasion (e.g. Ward, 2007, Giovanelli *et al.*, 2010, Hill *et al.*, 2012) as some statistical methods (e.g. Random Forest, Maxent, Stepwise GLM, GBM) are supposed to select automatically those variables with best discriminatory power.

2) Uncorrelated sets - V_{unc}: We built dendrograms where variables are clustered according to their pairwise correlations (Fig. 2). We define the distance matrix D as:

$$D = 1 - |R|$$

where R is the Pearson's correlation between all pairs of variables. Pearson's correlation is commonly used to select uncorrelated and complementarily informative variables (e.g. Graham, 2003, Dormann *et al.*, 2008, Braunisch *et al.*, 2013). The maximal number of variables resulting in a Pearson's correlation $\leq$ 0.7 was seven in NA and nine in EU, so we defined eight equidistant clusters of variables on each dendrogram, and randomly selected 1000 combinations of eight non-correlated variables, i.e. one in each cluster (Fig. 2).

3) Random sets - V_{ran}: We compared the V_{unc} strategy with eight variables selected randomly for 1000 replicates. Compared to V_{all}, this allows testing if reducing the number of variables and thus parameters decreases models complexities and the risk of over-parametrization (Rodder *et al.*, 2009). Additionally, this allows disentangling the possible effect to remove a certain amount of correlation when compared to V_{unc}.

4) State-of-the-art - V_{soa}: We selected eight variables that are commonly used in distribution models for plant species and which can be considered as a "state of the art" reference to model invasive plant species niches (Thuiller *et al.*, 2005, Broennimann *et al.*, 2007, Petitpierre *et al.*, 2012): Tmean, Tvar, Tcoldq, Twarmq, Pvar, Pwetq, Ma, Mvar (Table 2).

5) Stepwise hierarchical - V_{hie}: We used a stepwise hierarchical approach to select uncorrelated and important variables. Using statistical algorithms to pick up the most relevant variables is common in ecology (Mac Nally, 2002, Cutler *et al.*, 2007) and can be used in a hierarchical way (e.g. Roura-Pascual *et al.*, 2009). For each species, we built SDM models based on variables included in each cluster of the correlation dendrogram, respectively. Only the most important variable for the model within each cluster was retained such as in the end we got the eight most important and uncorrelated variables. When only one variable was included in a cluster (e.g. Twetq in EU), we automatically included it in the predictors set for the final model.

6) Most analog - V_{av}: We selected the most analog variables in the invaded range where we want to transfer the model. First, we measured the percentage of the invaded extent with non-analogue climate along each variable (Fig. 2). This is a similar approach as the Multivariate Environmental Similarity Surfaces (MESS, Elith *et al.*, 2010), but applied to each variable separately. Then we selected the eight variables with the lowest number of non-analogous sites without considering correlations between them. If more than eight variables had the highest analogy scores, we chose the variable with the highest distribution similarity between the native and the invaded range (corresponding to *P* in the sup. mat. of Elith *et al.*, 2010). To our knowledge, this approach have never been done despite several calls to take into such account factors analogy in variable selection (Rodder and Lotters, 2010, Mandle *et al.*, 2010, Soberón and Peterson, 2011)

7) Analog-uncorrelated - V_{avu}: Following a similar hierachical approach (as in 5), we selected the most transferable variables (as in 6) within each variables cluster of the correlation dendrogram, such as we get eight uncorrelated and analog variables.

8) Consensus - V_{con}: For each cluster of the correlation dendrogram, we ranked the variables in order to maximize both transferability and importance to select in a consensual way 8 uncorrelated, transferable and relevant variables. To do so, within each cluster, each variable gets two scores following their rank compared to the other variables within the same cluster: one depending their rank on climate analogy in the invaded range and one on variable importance determined following 5). Within each cluster, variables with the lowest averaged rank between climate analogy and importance were selected.

9) 8-axes PCA - V_{pc8}: We calibrated a principal component analysis on the 27 variables distributed in EU, NA and AU. We retained the first 8 components (i.e. axes). This approach is often used to reduce the number of parameters in the model and to decrease collinearity as components are orthogonal (Peterson *et al.*, 2007, Warren *et al.*, 2010, Warren *et al.*, 2008, Zhang and Zhang, 2011, Mellin *et al.*, 2010). Moreover, it has been shown to be the most accurate way to assess niche overlap (Broennimann *et al.*, 2012) and previous niche conservatism conclusions were based on this method (Petitpierre *et al.*, 2012) but did not imply SDMs calibrated on these components.

10) 8-axes PCA - V_{pc2}: Same as 9 but keeping only the first 2 components. The first two components explain 73% of the total climatic variation (Sup Fig. S1) while the first eight components explain 98%.

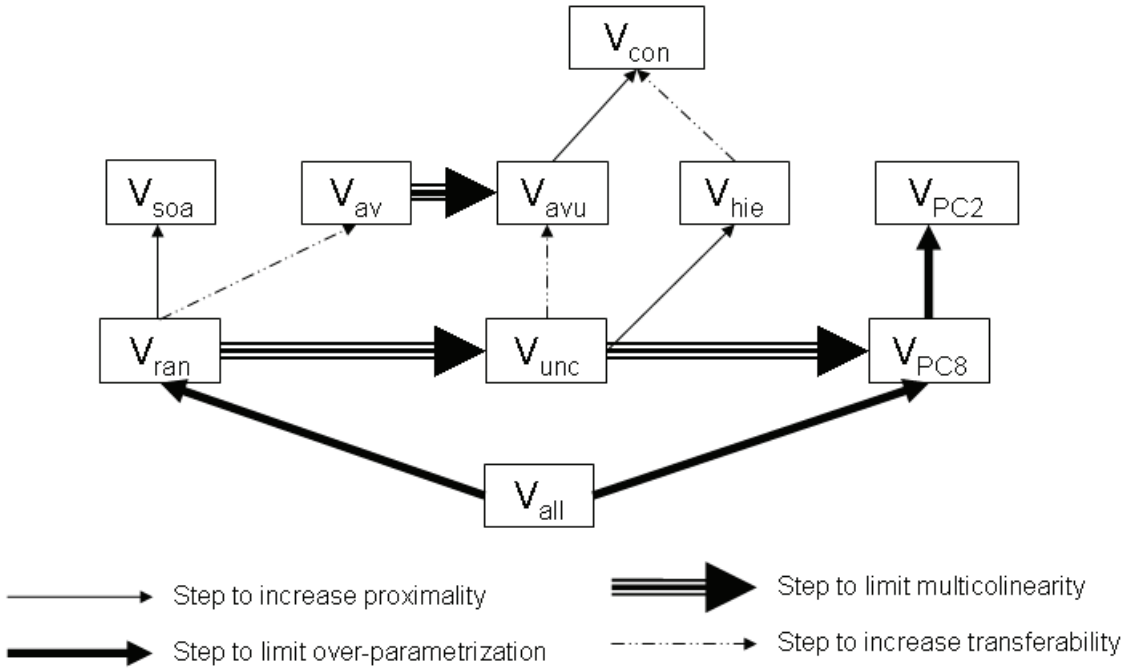


Figure 1: Scheme representing all variables datasets and their links in terms of proximality, over-parametrization, multicolinearity and climate analogy in the invaded range: exhaustive dataset (V_{all}), 1000 replicates of random 8 variables subsample (V_{ran}), 1000 replicates of variables with $R \leq 0.7$ of correlation (V_{unc}), "state-of-the-art" variables used to predict plant invasions at macroclimatic scale (V_{soa}), variables with more analogy in the invaded range (V_{av}), analog and uncorrlated variables (V_{avu}), uncorrelated and relevant variables (V_{hie}), uncorrelated, relevant and analog variables (V_{con}), 8 first components of PCA (V_{PC8}) and 2 first components of PCA (V_{PC2}). Note that all the possible links are not drawn, they rather reflect the spatial modeler's approach when one aims to work on the variable selection**.**

Modeling techniques

For each variables combination, we used three different modeling techniques: one regression technique (generalized linear models, GLM, McCullagh and Nelder, 1983), one machine-learning method based on decision trees (Generalized Boosted Regression Models, Friedman *et al.*, 2000) and one machine-learning methods based on maximum entropy (Maxent, Phillips *et al.*, 2006). Modeling was achieved using the R packages

BIOMOD for GLM and GBM (Thuiller *et al.*, 2009) and dismo for Maxent (Hijmans, 2010). Models were calibrated on 70% of the data in the native range and evaluated with the 30% of the remaining data in the native range and with the independent data in the invaded range(s). Variables contributions were estimated by assessing the impact on predictions of variables randomizations (see the documentation of BIOMOD for further details; Thuiller *et al.*, 2009). Predictions, evaluations and variables importance were averaged in a way to provide an ENSEMBLE model, resulting from a consensus of the 3 modeling techniques.

Evaluation

We used the True Skill Statistics (TSS) to evaluate model performances in the native range. TSS was recommended as an alternative to a commonly the commonly used Area Under the Curve of a Receiver Operating Characteristics (ROC AUC, Zweig and Campbell, 1993). As AUC, TSS also discriminates presences and absences correctly predicted but is less sensitive to species prevalence (Allouche *et al.*, 2006) and provides very similar results to the AUC (see e.g. Engler *et al.*, 2011 across 2632 species). Thus, TSS is appropriate in the native range where distributions are assumed to be closer their dispersal equilibrium. TSS varies between -1 and 1, with 0 for a random model, 1 for a perfect model and -1 for total counter predictions. In the invaded range, Petitpierre *et al.* (2012) showed that unfilling, i.e. environment not colonized in the invaded range but suitable in the native range is a widespread phenomena. It is thus problematic to use classical evaluation indices based on presences and absences like AUC or TSS. To evaluate models prediction accuracy in the invaded range, we thus used two metrics based on presence-only data. The Boyce index (B) measures how much predictions differ from the random expectation and varies between -1 and 1, with zero meaning not different than random. We used the same approach as Hirzel *et al.* (2006) to compute B along a moving window, relieving us from the constraint of fixing threshold to delimitate bins along continuous predictions. At last, we used sensitivity (Se), i.e. the percentage of presences correctly predicted by the model. It is particularly relevant in the case of invasive species where one aim to minimize commission error. Moreover, it is similar to the expansion index (E) proposed by Petitpierre et al. 2012 following the relationship:

$$Se = 1 - E$$

Note however that Se was calculated in the geographical space whereas E in the environmental space. In this manuscript we refer to bad, poor, fair, good, very good Se to values comprised between 0 - 0.5, 0.5 - 0.7, 0.7 - 0.8, 0.8 - 0.9 and 0.9 - 1 respectively. Se requires to be calculated from binary predictions maps, so we used the threshold providing the best TSS in the native range to classify the continuous predictions into presences and absences. Note that Se and B were also measured in the native range but not showed in the main manuscript (Fig. S2).

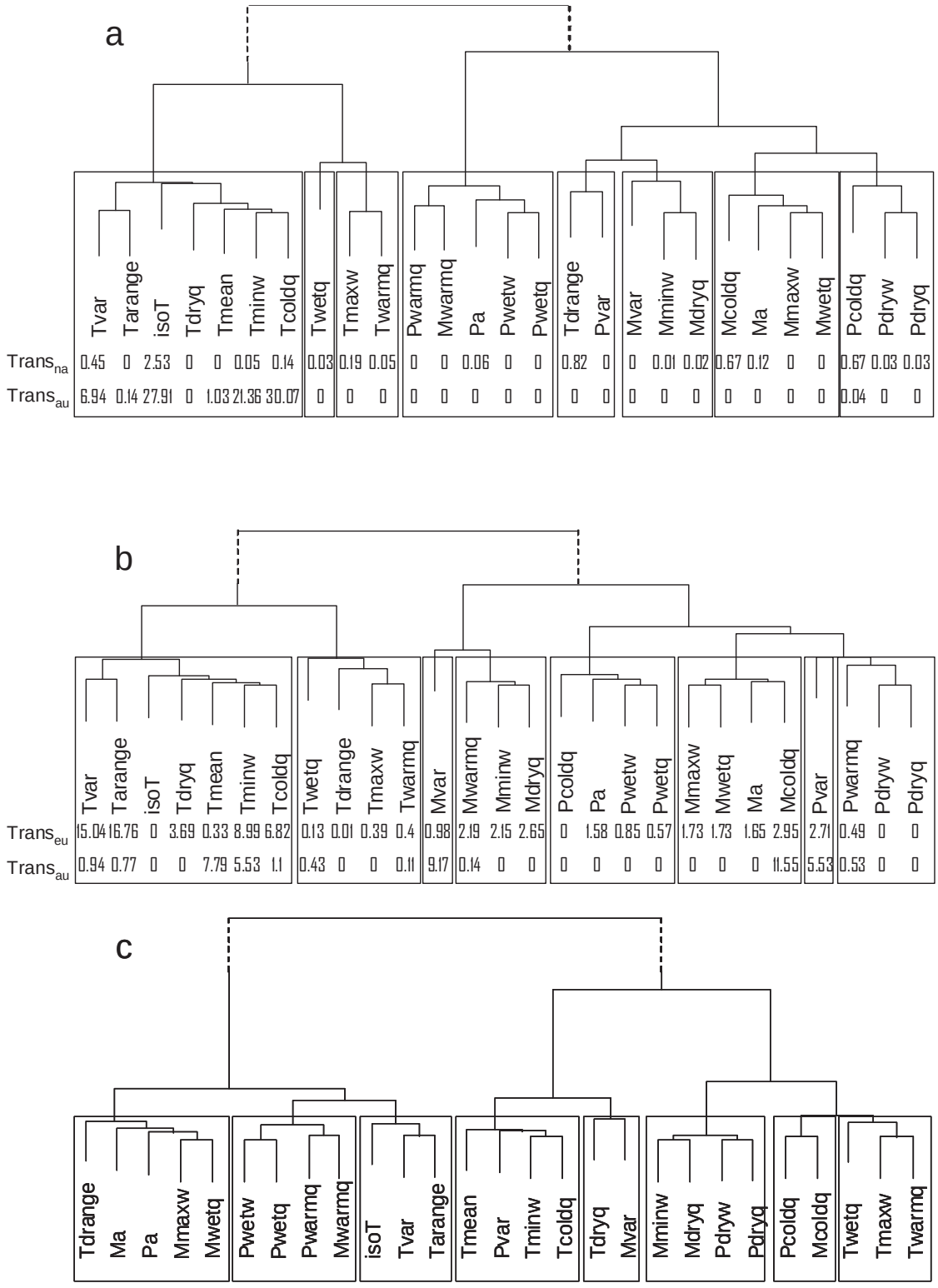


Figure 2: Dendrogram built on the Pearson's correlation between variables in EU (a) and NA (b). 8 Clusters sharing a similar average correlation have been delimited. Amount of non analogous climate (percentage of the invaded range showing different climate from the native range) is given for NA ($Trans_{na}$), AU ($Trans_{au}$) and EU ($Trans_{eu}$)

To determine the variables set allowing the best model transferability for each species, we chose the one maximizing (with the same weight) both B and Se in both Holarctic and Australian invaded ranges. A model with a high score on two different criteria, evaluated on 2 completely independent datasets is less likely to be obtained by chance. For species not present in AU, we considered then only the Holarctic invaded range.

To assess the impact of variable selection on projection under future climate for each species, we projected potential species distribution on the three ranges (EU, NA and AU) for each strategy and correlated them with projections of the best transferable model through space.

Relationship between variable species' traits, SDM transferability and variables contribution

We looked for association between species traits and SDM transferability in the invaded range for non-shifting species. For each species, ten traits were extracted from the Biolflor database (Kuhn *et al.*, 2004). Transferability was characterized using B and Se in the invaded range. We built GLM using a hierarchical approach to include parameters. First, we did a first traits selection by building a model for each traits category. For each category, we examined all traits combination retained the one providing the best BIC. Then these selected traits were pooled and all the possible combinations of the retained traits were examined to select the model with the best BIC. Models comparisons were done using the library MuMiN (Barton, 2012). We did this approach on the two most performing variables datasets, V_{soa} and V_{pc2}. Interactions between traits was not taken in account.

Additionally for each species, we correlated the variables contributions of the best model with 10 traits (Table 3). The relationship between the contributions of the 27 variables and the 10 traits was quantified through a MANOVA. The analysis was performed only with species that did not include principal components in their best model as principal components summarized the other variables and are thus non-independent.

RESULTS

Performance in the native range

In the native range, TSS ranged between 0.547 and 0.974, corresponding to good to excellent predictive power for most SDMs, except for *E. cilliatum* which had a lower but still fair TSS (Fig. 3, Table 11). Selection strategies had a significant effect on TSS (Kruskal-Wallis test, $P<0.001$), with V_{all} showing better performances than other strategies and V_{pc2} having lower TSS on the average. This trend also applied when SDM performances in the native range were evaluated with B and Se (sup Fig. 2) although the effect of the variable selection strategy was not significant when evaluated with B (Kruskall-test, $P = 0.119$).

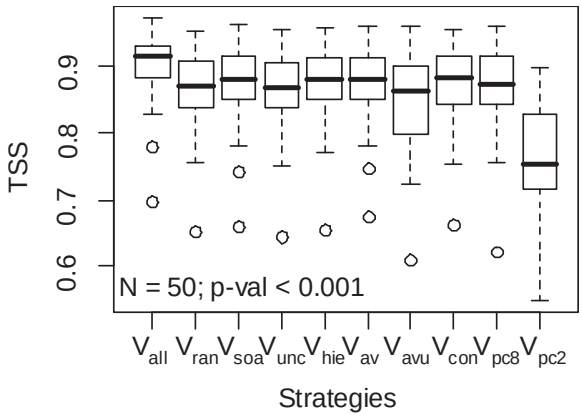


Figure 3: Average TSS for different variables selection strategies in the native range. N is the number of species included in the analysis and p-val is the significancy of a Kruskall-Wallis test

Performance in the invaded ranges

In their Holarctic invaded range, species showed B and Se values close to 1 on average but some species were poorly predicted with B below 0 and Se below 0.7 (Fig. 4, Table S2 and S3). Variables selection strategies had a significant effect on the average model performance for B and Se (Kruskall-test, P = 0.003 and < 0.001 respectively) but with different trends as when evaluated in the native range. V_{soa} and V_{pc2} had better evaluation scores on average for both B (0.83 ± 0.23 and 0.88 ± 0.08 respectively) and Se (0.85 ± 0.14 and 0.88 ± 0.11) and smaller variance in performances with less poorly predicted species than other strategies, especially V_{pc2} (*B. tectorum*, *E. esula*, *R. typhina*, *A. syriaca*, *A. novii-belgi* with V_{soa} and *E. esula* and *A*. *novii-belgi* with V_{pc2}, Fig. 4, Table S2 and S3). This translated into negative or weak correlations between evaluation on the native pseudo-independent data set and the invaded ranges (Fig. 5).

Models for species shifting their niche showed lower model performances on average in their Holarctic invaded range for both Se and B (t-test, P < 0.001) but strategies are affected differently by the magnitude of this decrease (Fig S4). Among seven species shifting in Holarctic, six species are poorly predicted with V_{soa} (*A. fruticosa*, *C. stoebe*, *C. scoparius*, *H. lanatus*, *T. dubium*, *H. tuberosus*) and 3 with V_{pc2} (*A. fruticosa*, *C. stoebe* and *C. scoparius*, Fig. 4, Table S2 and S3). Nonetheless, species shifting their niche in Holarctic did not present any significant decrease in model performance in AU (Fig. S4). When projected towards AU, performances of models for species shifting their niche in Holarctic are not significantly affected by the variables selection strategy although V_{pc2} provides the best results for both B (0.802 ± 0.178) and Se (0.867 ± 0.058) (Fig. S3).

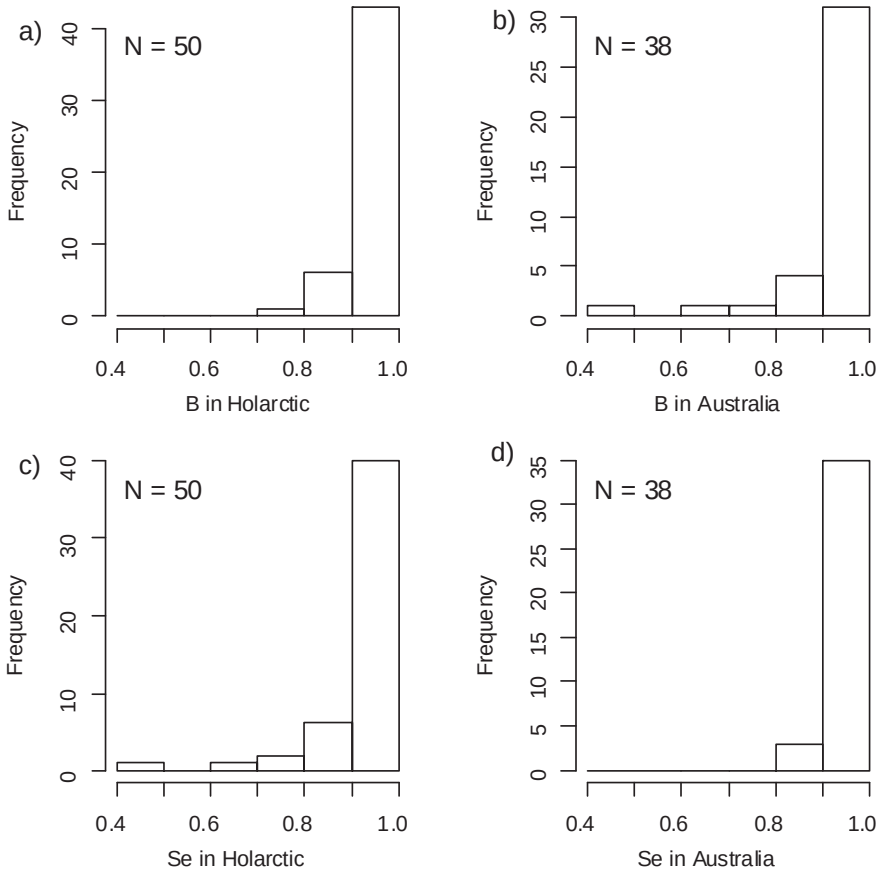


Figure 4: Evaluation of the model maximizing both Boyce index (B) and sensitivity (Se) in Holarctic and Australian (if available) ranges.

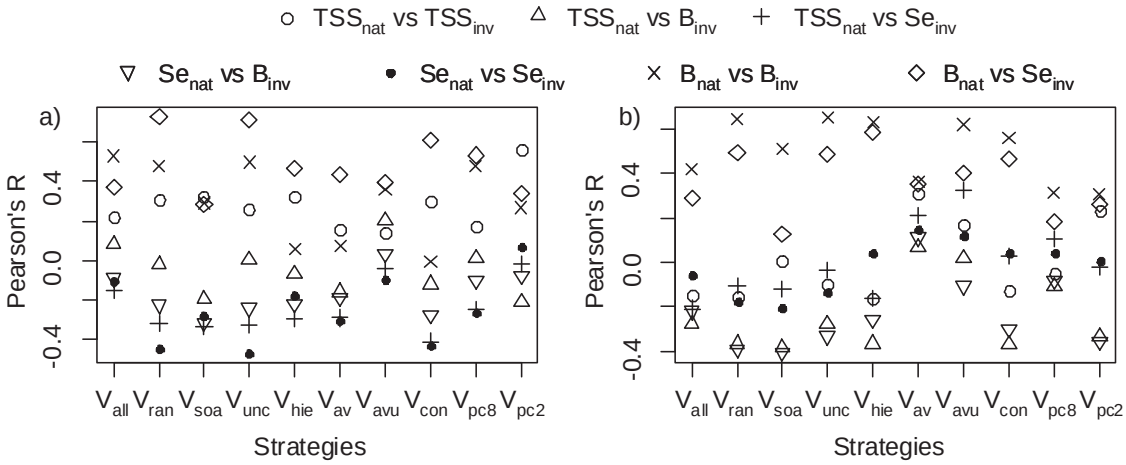


Figure 5: Correlation between SDM performances in the native range and in the invaded range in Holarctic (a) and in Australia (b) evaluated with True Skill Statistics (TSS), Boyce index (B) and Sensitivity (Se). This analysis only includes non-shifting species.

In the Australian invaded range, models show good performances on average but variables selection strategies do not show any significant effect (Fig. 4, Table S4 and S5). On the average, the V_{soa} and V_{pc2} strategy still shows the best performance with both B (0.77 ± 0.29 and 0.79 ± 0.24 respectively) and Se (0.89 ± 0.18 and 0.85 ± 0.19). V_{soa} also

includes the fewest poorly predicted species (*A. retroflexus* and *H. tuberosus*). Species shifting their niche in Australia show lower performance according to Se ($P < 0.001$) but not according to B ($P = 0.561$, Fig S5). Strategies based on transferable variables show the best performance for these species shifting their niche with V_{avu} providing the best B (0.851 ± 0.151) and V_{av} providing the best Se (0.815 ± 0.122) in the Australian invaded range, with only one species (*S. canadensis*) poorly predicted. In the Holarctic invaded range, species expanding their niche in Australia are not significantly affected by selection strategies: on average, V_{soa} and V_{pc2} are the best strategies with B (0.894 ± 0.135 and 0.861 ± 0.038 respectively) and Se (0.845 ± 0.101 and 0.841 ± 0.080 respectively, Fig. S4).

When focusing on the models maximizing both B and Se at the same time for every species, including all the replicates of the V_{ran} and V_{run}, we found only two species for which best models exhibited low predictions, even for most of the species shifting their niche in the invaded range (Fig. 6): *Centaurea stoebe* (Se = 0.478 in NA) and *Aster novii-belgi* (Se = 0.628 in EU and B = 0.442 in AU) (Fig.. 5, Table 1). We observed that best models are always provided by the random (V_{ran}: 36 species), the random/uncorrelated (V_{unc}:11 species) or PCA2 (V_{PC2}:3 species) strategies (table 1).

GLM based on traits explained between 10% and 38% of the variation of transferability (sup Fig. 5). Different traits were retained among strategies, evaluation metrics and location of the invaded ranges (Table 4, Fig. S5, Table S6). Among the most conserved trends between strategies and evaluators, we observed a positive correlation between native niche breadth and transferability in Holarctic and a negative effect on transferability of the North-American native for species introduced in Australia (Table 4, Table S6). More marginally, we observed possible positive effect of decreasing ploidy level, selfing reproduction and a negative effect for species able of vegetative reproduction (Table 4 and Table S6).

Projections under climate change

Projections of SDM under climate change scenario are also affected by variables selection. When we compare projections of the best model with projections provided by the other variables selection strategies, Pearson's correlations are high and vary between 0.725 ± 0.2 (selection based on uncorrelated variables and transferability towards AU, V_{avuau}) and 0.874 ± 0.09 (selection based on expert knowledge, V_{soa}) (Fig. 6) and variable selection strategies have a significant effect on this correlation (Kruskall-test, $P < 0.001$). The differences between projections can result in important divergences between potential distributions of the species under climate warming (e.g. Fig. 7). Variable selection strategies have no significant effect on species shifting shifting their niche, although species shifting their niche in Holarctic keep the same trend as non-shifting species (Fig. S9).

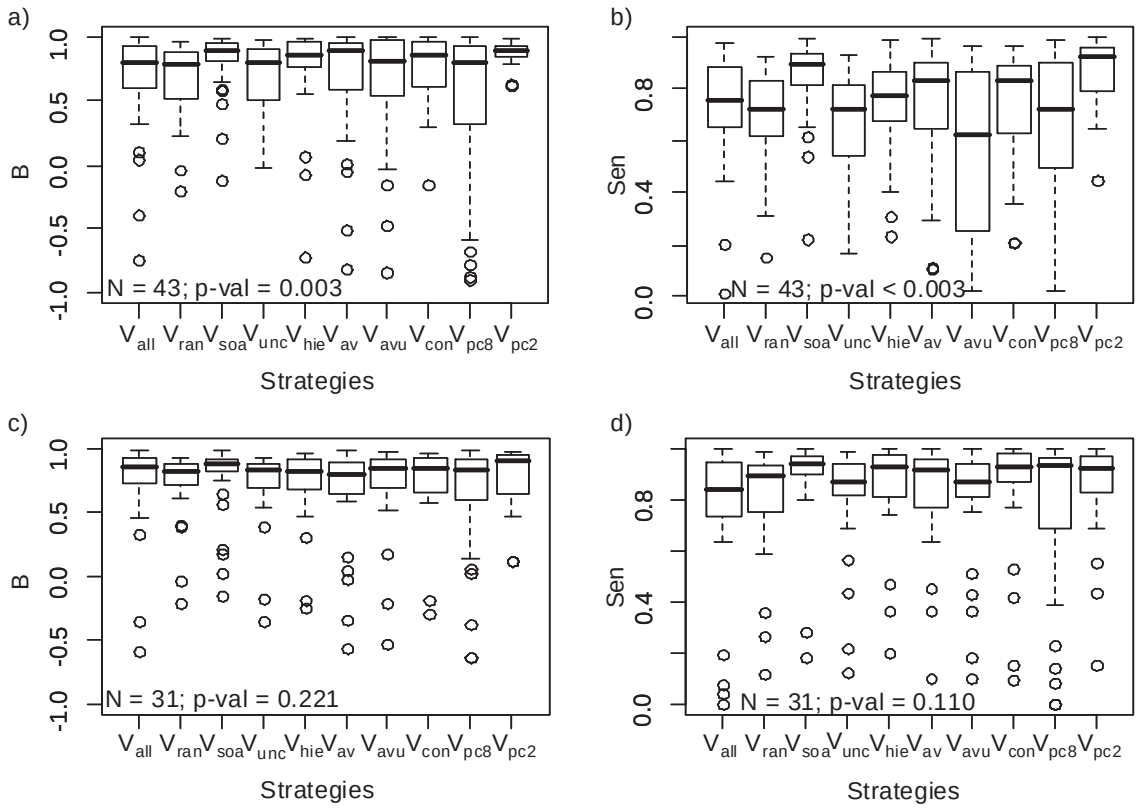


Figure 6: Evaluation in the invaded range following different variables strategies. Models are evaluated with Boyce index (a and c) and sensitivity (b and d) when projected onto their reciprocal Holarctic invaded range (EU or NA, a and b) and in AU (c and d). Number of species included in the analysis (N) and and p-value of a Kruskall-Wallis test is provided in each case.

Variable importance and species' traits

In the best models, the 8 most frequently included variables are, in ranking order, Pvar, Mdryq, Mvar, Twarmq, Tmean, Mwarmq, Pdryq, Mmaxw. Tdrange is included in only 3 best models. The two first principal components provided the best models for three species (Fig. 8). The only trait presenting a possible effect on variable contributions is species origin, with a significant effect when all species are pooled together (MANOVA pillaj-test, P = 0.030 Table 3). Once they are included, temperature variables have higher contribution than the variables in other categories. This trend is also confirmed by the more important contributions of the second component of the PCA, corresponding to temperature variables, when PCA provides the best model (Fig 9) and is also visible for the analysis of species shifting their niche (Fig. S10 - S11).

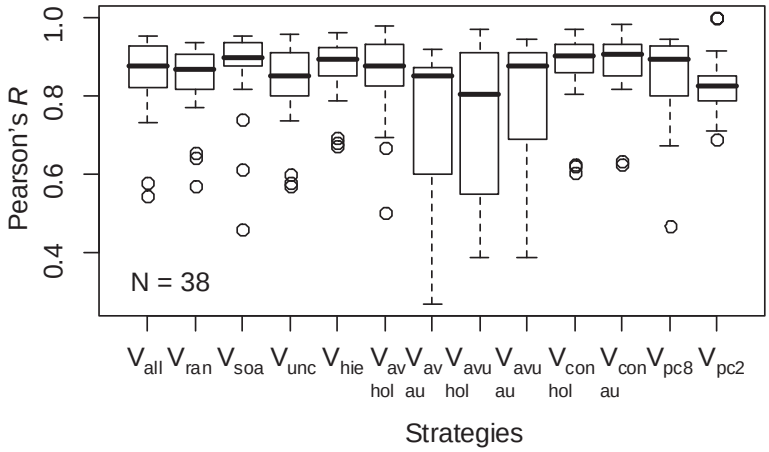


Figure 7: Pearson's correlation between best models projections under future climate in EU, NA and AU and projections provided by variables selection strategies. Note that species shifting their niche either in Holarctic or Australian ranges were not included in this analysis.

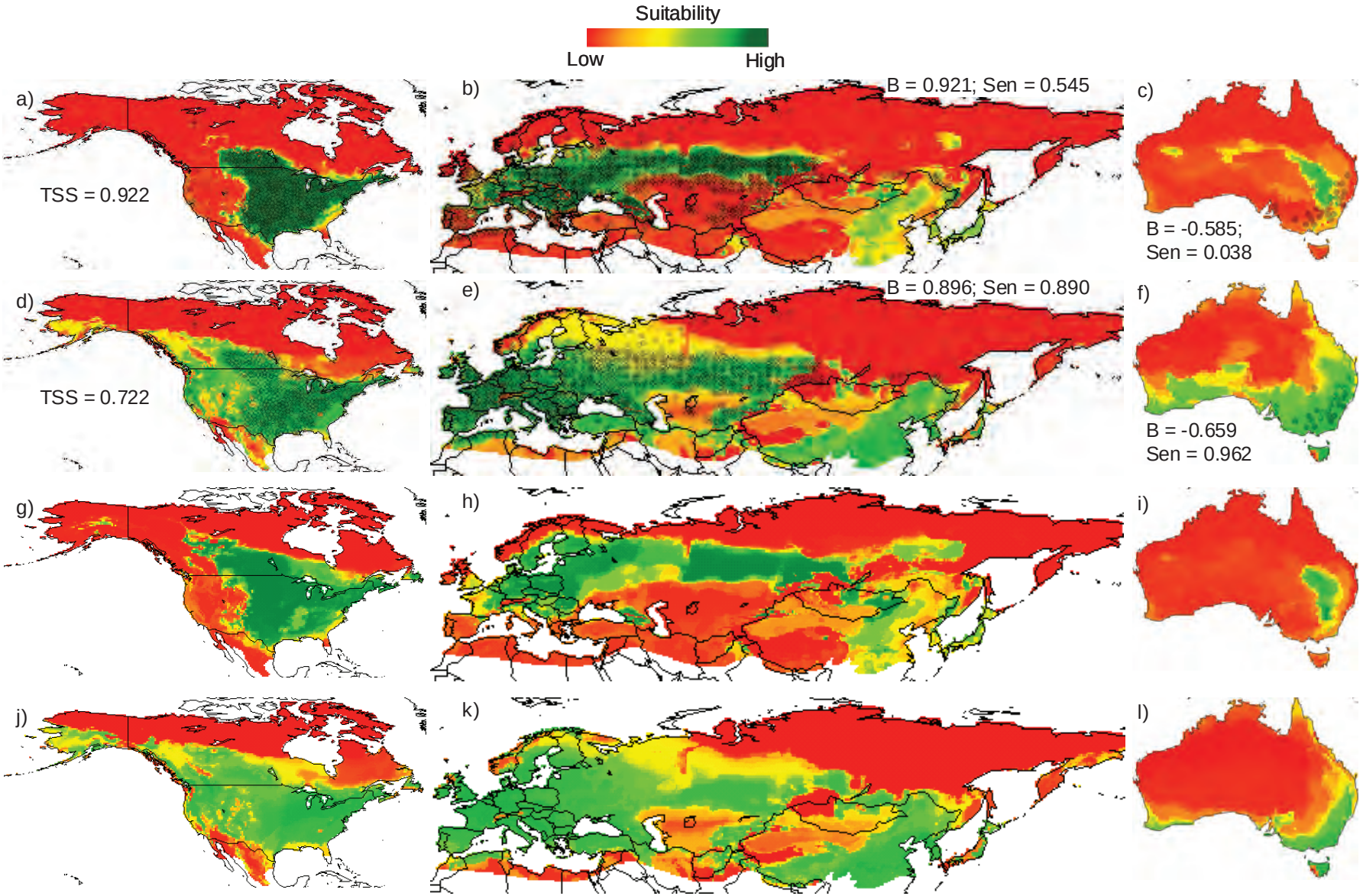


Figure 8: projections of the best model (V_{pc2}, a,b,c,g,h,i) and the model including all variables (V_{all}, d,e,f,j,k,l) under current (a-f) and future (g-l) climatic conditions for A. retroflexus. Actual distribution of the species is represented by dots.

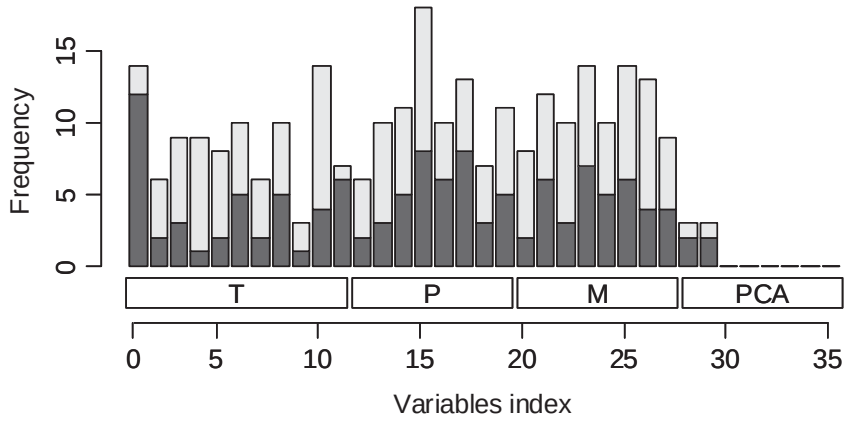


Figure 9: Number of times variables are included in the best model of non-shifting species (N = 38). Black and grey represent EU and NA species respectively. The variables index is the same as table 2. T, P, M and PCA represent temperature, precipitation, moisture and principal component variables

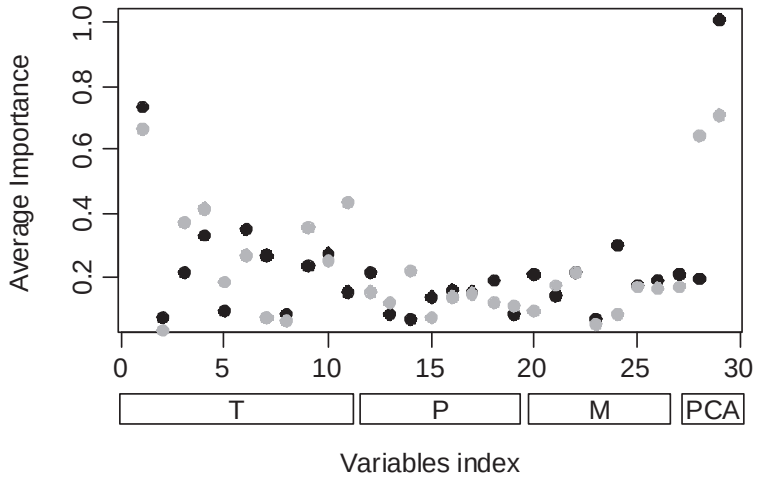


Figure 10: average variables importance once they are included in the best model for non-shifting species (N =38). Black and grey represent EU and NA species respectively. T, P, M and PC represent temperature, precipitation, moisture and principal component variables. The variables index is the same as Table 2

DISCUSSION

Our results show that SDM can be transferred in space for 48 species out of 50 but that variables selection significantly impacts model performances in the invaded range. On average, the standard set of variables (V_{soa}) and a reduced and orthogonalized set (V_{pc2}) provide the best SDM transferability into Holarctic. However, when projecting into a more differentiated environment such as Australia, although V_{soa} remains robust, the transferability of specific predictors may be taken in account (as in V_{tra} or V_{tu}), as transferable variables provide better SDMs for Australian species shifting their niche in Australia. Overall, this advocates for proximal datasets and simple models.

Evaluation dataset

The asymmetry between performances in the native and invaded ranges with V_{all} and V_{pc2} demonstrates that excellent performances on pseudo-independent data do not imply a good transferability. Spatial autocorrelation and over-parametrization can explain this apparent paradox. The usual approach consisting in using a subsample of the calibration area as an independent dataset for models evaluation may be biased by spatial correlation with the calibration dataset (Araujo *et al.*, 2005, McPherson and Jetz, 2007). Thus, SDMs overfitting the calibration dataset would also predict the pseudo independent dataset without full test of predictive power. To minimize this spatial autocorrelation problem, increasing the ratio of independent data in the split-sampling, including a spatial autocorrelation term or disaggregating the calibration dataset (Dormann, 2007) can be an alternative. Note however that spatial disaggregation (also known as occurrences thinning) may not be relevant if there is no sampling bias in the calibration dataset (Verbruggen *et al.*, 2013).

Evaluation metric

The way models are evaluated can also impact the assessment of SDM transferability. In the native calibration dataset, we assumed that the realized niche is at equilibrium with the environment, thus evaluating SDMs ability to discriminate presences from absences, attributing the same weight to both. Such an evaluation approach is particularly relevant to depict the realized niche optimum but occurrences located at the margin of the niche, where continuous models suitabilities are lower, tend to be considered as absences when evaluated using a binarizing threshold. The realized niche is equal to or smaller than the fundamental niche except in the case of biotic facilitations. Thus models closer to the fundamental niche, predicting a wider potential distribution, also allow correctly predicting occurrences at the margin of the realized niche. This approach may be more conservative than optimizing for both sensitivity and specificity, even though the rate of false positive (Type I error) is increased (Vale *et al.*, 2013). False presences can be simply due to non-equilibrium in the geographical range or to competitive interactions. In the case of biological invasions and rapid climate change, biotic interactions are subject to change and modeling closer to the fundamental niche provides more conservative predictions. To do that, adding presence-only evaluators such as sensitivity, Boyce, POC-plot (Phillips and Elith, 2010), or varying the sampling and the prevalence of the background pseudo-absences (Barbet-Massin *et al.*, 2012, Meynard *et al.*, 2013) may be helpful to build more transferable models.

Selecting the right variables to transfer models

Among the 10 variables selection strategies that we tested, V_{soa} and V_{pc2} performs the best on average, but the best model is provided by a randomly selected data set (V_{ran} and V_{unc}) for 47 species and V_{pc2} for 3 species. By screening a large set of random variables

combinations, we increase the chance to find singular combinations which predict the invaded range. The fact that variables sets providing the best model are mostly randomly selected supports there is no strategies giving the best model for all species (Williams *et al.*, 2012); the best variables set is proper to each species. Note that the best model we found is not the best of all possible variables combination but the best among a sample of 2000 different variables combinations.

Niche conservatism between native and invaded range is a pivotal assumption to project SDM through space and time (Pearman *et al.*, 2008*a*, Peterson, 2011). As variables selection impacts SDM transferability, the degree of niche conservatism may vary from a variable to another. Such shifts can be independent from ecological or evolutionary processes affecting species fitness and simply result from non-proximal variables confounded with important variables for the delimitation of species distribution or from climatic non-analogy in the native range. It is thus important to understand the nature of apparent niche shifts along variables used to depict the species realized niche. We tested if the expansion rate of species shifting their niche was lower than with the set of variables providing the best model than with the "state-of-the-art" variables set (V_{soa}) analogous to the one used in Petitpierre *et al.* (2012) and found no significant differences for Holarctic shifting species (t-test, P = 0.568) but a significant difference for species shifting their niche in Australia (t-test p-val <0.001). Thus in the Holarctic, the realized niches are not significantly more conserved with the set of variables providing the best model whereas in AU, expansion is lowered for species shifting their niche in AU. Note that for some other species presenting less than 10% of niche expansion in Petitpierre *et al.* 2012, E is increased when using the variables set providing the best model (e.g. *Holcus lanatus*, *Verbascum thaspum*, *Aster novi-belgii*, *Epilobium ciliatum*, *Helianthus tuberosus*).

Expansion may occur at one side of each gradient instead of both. Indeed, the realized niche can be more labile at one or another extremity of the gradient and it has been shown that the most stressful extremity of the gradient is more predictable by SDMs (Normand *et al.*, 2009, Maiorano *et al.*, 2012, Pottier *et al.*, 2013, Araújo *et al.*, 2013) as it is closer to the realized niche. This can be seen for moisture and isothermality for *Cytisus scoparius*, *Holcus lanatus*, *Trifolium dubium* (Fig. S10).

It is recommended to use proximal variables, known to have direct impact on species physiology and fitness, to predict potential species distribution (Kearney and Porter, 2009, Morin and Thuiller, 2009, Buckley *et al.*, 2010). Predictions based on the fundamental niche assessed through physiological experiments would be more conservative, even if the fundamental niche may also be subject to changes, but requires evolutionary adaptations which take times to develop (Whitney and Gabler, 2008). Variables set providing more transferable models can be used as a first step to investigate factors affecting species fundamental niche as experimental data on species physiological response to environmental gradient are rarely available.

A second non-biological cause of niche shift in the realized niche can be predictions projected outside the climatic range used in the calibration, i.e. to novel, non-analog

conditions (Fitzpatrick and Hargrove, 2009, Mandle *et al.*, 2010, Peterson, 2011). Extrapolating complex SDM may lead to spurious response as there is no guarantee that interactions remain constant and no other inflexion points are added in the novel conditions characterizing the invaded range. Strategies based on climate analogy do not show better performances for non shifting species even in AU where climate is quite differentiated (Fig. S12) but better performances for species shifting their niche in AU, suggesting that nature of this shift in AU could be linked climatic non-analogy in the native range. However, observed expansion does not occur within non-analog climates (Fig S10 and S11) making this suggestion highly speculative and difficult to interpret.

It is interesting to note that collinearity does not show any significant negative effect on predictions (e.g. when V_{unc} is compared to V_{ran}). Using Pearson's correlation to assess variables similarity is very common but can be subject to criticism. First, the threshold (here in this study 0.7) is empirically fixed and does not rely on any statistical demonstration or simulation. Second, this approach is biased when non-linear relationships exist among predictors. A possible alternative to evaluate the similarity among predictor is the use of matrix based on indices such as Gower metric (Williams *et al.*, 2012). However, both the validity of the correlation threshold and Gower matrices require to be rigorously assessed. Independent dataset such as invasive distribution can be useful for such purpose.

In the case where specific information about proximal predictors affecting species fitness or in cases where many species niche are modeled and comparable variables datasets are more desirable, variables set such as V_{soa} or V_{pc2} are good alternatives. On average, V_{soa} performs well on every range and is highly correlated with projections of the best model under climatic scenarios. Thus it seems that although including extremely few parameters and thus very simple models, V_{pc2} performs well except with shifting species in AU and generally presents less variance and poorly predicted species. Note that SDM calibrated on principal components may be more problematic to interpret and are less correlated with projections of the best model with climate change. Climatic scenarios may involve changes in correlation structure between parameters and thus lead to spurious effect when projected onto the PCA environmental space. Note also that in our case, most of the climatic variation is included in the two first components, thus adding more components seems to add noisier and ecologically irrelevant parameters to the SDMs. However, if less variation was summarized in the two first components, adding more components and thus more information could increase model transferability.

To summary, if no physiological information is available concerning species of interest, V_{pc2} could be a relevant alternative by providing simple models when projecting onto predictors keeping the same correlation structure. V_{soa} may be more desirable if you are interested in ecological interpretation or in projection towards climatic scenarios where predictors may have different correlation structures. Eventually, when the projection range takes place in a highly differentiated environment, strategies based on variables analogy (such as V_{av} or V_{avu}) can be considered. Of course, consensus prediction based on

different variables set allows showing all possible projections with their convergences and divergences.

Species traits and transferability

It is difficult to predict species invasiveness based on traits as most of the attempts have been idiosyncratic (Williamson, 1999, Mack *et al.*, 2000, Colautti *et al.*, 2006) and seems context dependent (Daehler, 2003, Kueffer *et al.*, 2013), although potential perspectives through the combination of SDMs and traits analyses (Moles *et al.*, 2008). Note that we removed species showing major niche changes in their invaded range and that these remaining species are not expected to be affected by severe niche changes at this macroclimatic scale. However, we found two consistent traits associated with transferability using V_{soa} and V_{pc2}. Species with wider native niche breadth are more transferable in their Holarctic invaded range and species originating from NA are less transferable in AU. Smaller native niche may be the symptom of dispersal limitation in the native range and thus a possible non-environmental equilibrium. Additionally, species with a wider niche can colonize more environments and thus decrease the probability of being introduced outside of its native niche. Disentangling this statistical effect from possible non-environmental equilibrium in the native range is tricky and requires further investigations.

Species native origins has already been associated with different niche dynamic patterns in the invaded ranges, with NA species showing less niche expansion and more unfilling in AU (Petitpierre *et al.*, 2012). Such pattern should translate into better transferability of NA species, at least when evaluated with Se (because inversely proportional to niche expansion). Indeed, transferability assessed with B index is most severely affected by this origin factor. On the contrary, we can see more unpredicted EU species when their potential distribution is predicted with the best model (Fig. S12).

Other factors affecting transferability only on one way to evaluate transferability, one invaded range and one strategy appeals for further investigation to determine if they play actual roles in SDM transferability in general. However, it has been shown that polyploidy provides a better evolutive potential, making species more susceptible to adapt more rapidly to novel conditions with potential fundamental niche shifts. On the contrary, species able to self pollination are more exposed to genetic bottleneck and thus have less probability to evolve.

We did not find any significant influences between species traits and variables contributions in their best model. The only trait showing a trend is species origins again. For variables contributions, our analysis shows that the difference between EU and NA species occurs in the inclusion rate of variables rather than their contribution once they are included. Comparably to Pellissier *et al.* (2013) and Randin *et al.*, 2013), thermal variables also show more importance to depict species potential distributions. This can be seen as good news to forecast species distribution under climate change scenarios. Indeed,

scenarios about precipitation are more uncertain than temperature scenarios (Bosshard *et al.*, 2011).

Conclusion

Our study shows that among 50 Holarctic plant invaders, it is possible to predict the invaded range for 48 species from their native range. It is thus possible to project SDMs through space at a coarse continental scale and would support the transferability of SDMs through time when rapid climate changes scenarios are involved. The variables set used to build SDM significantly impacts transferability. Otherwise, strategies based on fewer parameters or traditional datasets used to build SDMs are the best alternative to build predictive SDM on average.

Acknowledgments

The computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics. We thanks Christophe Randin for previous feed-back on the manuscript. AG, OB and BP received main support from the National Center for Competence in Research « Plant Survival ». CD and CK received support from the USDA National Institute of Food and Agriculture, Biology of Weedy and Invasive Species Program Grant no. 2006-35320-17360.

Table 1: Evaluation of best models for each species with True Skill Statistics in the native range (TSS), Boyce index and Sensitivity in the Hoarctic and Australian invaded range (B_{hol}, Sen_{hol}, B_{au}, Sen_{au} respectively). [1)] and [2)] indicate species shifting in Holarctic and Australia respectively. Included variables and their importance in the model and species origin are also provided.

Sp name	Sp origin	Strategy	TSS	B hol	Se hol	B au	Se au	Variables
Alliaria petiolata	NA	Vran	0.822	0.987	0.938	0.895	1	Tmean(0.94);Pa(0.142);Pwetw(0.021);Pdryw(0.101);Pwetq(0.027);Pdryq(0.109);Ma(0.228);Mdryq(0.117);
Amaranthus retroflexus	EU	Vru	0.856	0.974	0.997	NA	NA	Tdrange(0.022);Twetq(0.05);Tdryq(0.529);Twarmq(0.063);Pwetw(0.044);Pcoldq(0.391);Mwetq(0.019);Mdryq(0.164);
Ambrosia artemisiifolia	NA	Vpc2	0.72	0.896	0.89	0.659	0.962	PCA1(0.232);PCA2(1.026);
Amorpha fruticosa[1]	NA	Vru	0.883	0.989	0.866	0.933	1	isoT(0.166);Twarmq(0.773);Pvar(0.02);Pwetq(0.031);Pdryq(0.051);Mvar(0.048);Mdryq(0.141);Mcoldq(0.014);
Anagallis arvensis	NA	Vran	0.754	0.803	0.839	NA	NA	Tmean(0.409);Tvar(0.345);Tarange(0.118);Twarmq(0.591);Tcoldq(0.212);Pwarmq(0.16);Pcoldq(0.181);Mwarmq(0.153);
Anthoxanthum odoratum	EU	Vru	0.89	0.96	1	0.993	1	isoT(0.717);Twetq(0.016);Twarmq(0.054);Pvar(0.052);Pdryq(0.228);Mmaxw(0.084);Mdryq(0.148);Mwarmq(0.101);
Arabidopsis Thaliana	EU	Vran	0.891	0.97	0.997	0.967	0.994	Tvar(0.599);Tminw(0.34);Tarange(0.145);Tcoldq(0.258);Pdryw(0.14);Mminw(0.093);Mwetq(0.067);Mcoldq(0.211);
Bromus sterilis[1]	EU	Vru	0.852	0.987	0.993	0.89	1	isoT(0.16);Twetq(0.029);Twarmq(0.207);Pdryw(0.061);Pvar(0.294);Ma(0.112);Mvar(0.03);Mwarmq(0.085);
Bromus tectorum	NA	Vran	0.914	0.936	0.988	NA	NA	Tmean(0.876);isoT(0.162);Tarange(0.105);Pdryw(0.101);Pdryq(0.181);Mmaxw(0.095);Mminw(0.234);Mcoldq(0.061);
Carduus nutans	NA	Vran	0.9	0.703	0.629	0.442	0.85	Tmean(0.251);Tdrange(0.067);Tmaxw(0.205);Twarmq(0.486);Pwetw(0.186);Pwetq(0.376);Ma(0.398);Mwetq(0.552);
Centaurea stoebe[1]	NA	Vran	0.847	0.94	0.965	NA	NA	Tmean(0.664);Tmaxw(0.298);Pwetw(0.009);Pwarmq(0.027);Pcoldq(0.062);Mmaxw(0.3);Mvar(0.041);Mwetq(0.107);
Cirsium vulgare[2]	EU	Vran	0.932	0.895	0.933	0.843	0.971	Tmaxw(0.176);Twarmq(0.23);Tcoldq(0.981);Pwetq(0.001);Pwarmq(0.078);Mmaxw(0.138);Mvar(0.11);Mwetq(0.28);
Conyza canadensis	EU	Vpc2	0.827	0.927	0.795	0.926	1	PCA1(0.159);PCA2(0.988);
Cytisus scoparius[1]	EU	Vru	0.84	0.922	0.985	0.912	1	Tvar(0.429);Tmaxw(0.122);Twetq(0.085);Pvar(0.06);Pdryq(0.183);Mmaxw(0.062);Mdryq(0.246);Mwarmq(0.269);
Dactylis glomerata	EU	Vran	0.896	0.964	0.478	NA	NA	Tmean(0.509);isoT(0.219);Tmaxw(0.084);Twetq(0.071);Pcoldq(0.163);Ma(0.131);Mminw(0.156);Mvar(0.11);
Echinocystis lobata	EU	Vran	0.89	0.98	0.904	0.972	0.957	Tvar(0.286);Twetq(0.029);Twarmq(0.223);Pdryw(0.019);Pwetq(0.076);Pdryq(0.333);Pcoldq(0.09);Mvar(0.004);
Erigeron annuus	NA	Vran	0.936	0.887	0.954	0.912	1	Tmean(0.932);isoT(0.202);Tminw(0.133);Tcoldq(0.156);Pdryw(0.021);Pvar(0.014);Mminw(0.031);Mwetq(0.054);
Erodium cicutarium	EU	Vran	0.926	0.877	0.935	0.953	1	isoT(0.658);Twarmq(0.275);Pwetq(0.148);Pwarmq(0.208);Mmaxw(0.144);Mvar(0.291);Mwetq(0.536);Mdryq(0.237);
Euphorbia esula	EU	Vran	0.839	0.988	0.968	0.962	0.995	isoT(0.34);Tmaxw(0.062);Twarmq(0.247);Pvar(0.301);Pwetq(0.183);Pwarmq(0.092);Mminw(0.259);Mdryq(0.143);
Holcus lanatus[1]	NA	Vru	0.886	0.991	0.952	NA	NA	Tmean(0.714);Twetq(0.14);Pvar(0.004);Pdryq(0.043);Pcoldq(0.141);Mmaxw(0.142);Mvar(0.091);Mdryq(0.306);
Hypochaeris radicata[2]	NA	Vran	0.658	0.992	0.973	0.901	0.964	Tmean(0.245);Tvar(0.184);Tarange(0.006);Twarmq(0.276);Tcoldq(0.742);Pdryw(0.301);Pvar(0.05);Pdryq(0.305);
Juncus tenuis	NA	Vran	0.953	0.976	0.957	NA	NA	Tmean(0.785);Tminw(0.155);Twetq(0.02);Tcoldq(0.177);Pwetq(0.088);Mwetq(0.122);Mdryq(0.19);Mwarmq(0.059);
Linaria vulgaris[2]	EU	Vran	0.873	0.995	0.909	0.974	0.966	Tmean(0.797);Tmaxw(0.081);Twarmq(0.184);Pa(0.292);Pvar(0.311);Ma(0.112);Mvar(0.139);Mdryq(0.149);
Lythrum salicaria	EU	Vran	0.843	0.922	0.792	NA	NA	Tmean(0.687);Pa(0.089);Pwetq(0.141);Pdryq(0.1);Pwarmq(0.027);Pcoldq(0.027);Mminw(0.21);Mvar(0.044);
Medicago lupulina	NA	Vran	0.909	0.932	0.885	0.727	0.909	Tmean(1.057);Tminw(0.2);Tcoldq(0.268);Pa(0.083);Mminw(0.186);Mdryq(0.002);Mwarmq(0.031);Mcoldq(0.06);
Melilotus albus[2]	EU	Vran	0.921	0.957	0.895	0.944	0.995	isoT(0.552);Tmaxw(0.072);Pdryw(0.1);Pvar(0.208);Pwetq(0.114);Ma(0.192);Mmaxw(0.119);Mcoldq(0.33);
Phytolacca americana	EU	Vran	0.924	0.936	0.973	0.977	0.991	isoT(0.62);Tmaxw(0.097);Pwetw(0.09);Pdryq(0.268);Pcoldq(0.198);Ma(0.069);Mwarmq(0.093);Mcoldq(0.298);
Plantago lanceolata	NA	Vru	0.873	0.971	0.963	0.907	1	Tmean(0.733);Twetq(0.093);Pvar(0.066);Pdryq(0.082);Pcoldq(0.068);Mmaxw(0.163);Mvar(0.069);Mdryq(0.114);
Plantago major	EU	Vran	0.843	0.985	0.943	0.893	1	Tmaxw(0.122);Tdryq(0.327);Pvar(0.094);Pdryq(0.368);Mmaxw(0.239);Mminw(0.419);Mvar(0.114);Mwetq(0.35);
Poa annua	EU	Vran	0.877	0.958	0.938	0.947	0.936	Tvar(0.48);Pwetw(0.054);Pwarmq(0.294);Mmaxw(0.096);Mminw(0.26);Mdryq(0.156);Mwarmq(0.274);Mcoldq(0.609);
Potentilla recta	EU	Vran	0.759	0.989	0.968	0.955	0.984	Tmaxw(0.135);Tminw(0.38);Twarmq(0.235);Pwetw(0.271);Pdryw(0.152);Pvar(0.144);Mvar(0.056);Mwetq(0.14);
Prunus serotina	EU	Vran	0.772	0.998	0.96	0.954	0.963	Tarange(0.202);Tdryq(0.299);Pdryw(0.436);Pwetq(0.355);Ma(0.34);Mvar(0.112);Mdryq(0.189);Mcoldq(0.562);
Rhus typhina	NA	Vran	0.861	0.946	0.982	0.921	1	Tminw(0.442);Tdryq(0.191);Tcoldq(0.663);Pdryq(0.092);Pwarmq(0.251);Mmaxw(0.227);Mvar(0.071);Mwarmq(0.14);
Robinia pseudoacacia	EU	Vran	0.854	0.985	0.918	0.981	0.985	Tminw(0.843);Tarange(0.088);Pvar(0.242);Pwetq(0.097);Pcoldq(0.084);Mminw(0.171);Mvar(0.023);Mdryq(0.154);
Rumex acetosella	EU	Vru	0.783	0.971	0.971	0.98	0.985	isoT(0.245);Twetq(0.021);Twarmq(0.267);Pvar(0.094);Pwetq(0.206);Pcoldq(0.067);Ma(0.125);Mdryq(0.073);
Solidago canadensis[2]	EU	Vran	0.81	0.961	0.934	0.931	0.977	Tmaxw(0.047);Tminw(0.571);Tarange(0.125);Twarmq(0.277);Pdryq(0.336);Pcoldq(0.072);Mmaxw(0.112);Mvar(0.022);
Solidago gigantean	EU	Vran	0.853	0.996	0.993	0.917	1	Tdrange(0.131);Tvar(0.182);Tarange(0.579);Tdryq(0.238);Pdryw(0.127);Mminw(0.407);Mwarmq(0.341);Mcoldq(0.145);
Sonchus oleraceus[2]	NA	Vru	0.936	0.989	0.988	NA	NA	Tmean(0.642);Twarmq(0.256);Pdryw(0.027);Pvar(0.016);Pcoldq(0.053);Mmaxw(0.286);Mminw(0.114);Mvar(0.114);
Trifolium arvensis	NA	Vpc2	0.888	0.916	1	NA	NA	PCA1(0.648);PCA2(0.712);
Trifolium dubium[12]	NA	Vru	0.795	0.963	0.962	0.981	0.956	Tmean(0.744);Tdrange(0.035);Pvar(0.02);Pwetq(0.063);Pdryq(0.094);Mvar(0.014);Mwarmq(0.071);Mcoldq(0.028);
Trifolium repens	NA	Vran	0.815	0.987	0.98	NA	NA	Tmean(0.904);Tminw(0.1);Twetq(0.026);Tcoldq(0.018);Pwetq(0.049);Mwetq(0.046);Mdryq(0.19);Mwarmq(0.059);
Verbascum thapsus	EU	Vran	0.85	0.964	0.952	0.97	0.95	Tdrange(0.046);Tminw(0.312);Twarmq(0.157);Pwetw(0.169);Mmaxw(0.221);Mminw(0.175);Mwetq(0.101);Mcoldq(0.331);
Vicia sativa	NA	Vru	0.827	0.865	0.933	0.911	0.865	Tdryq(0.144);Twarmq(0.675);Pdryw(0.054);Pvar(0.117);Pcoldq(0.048);Mvar(0.118);Mwetq(0.173);Mdryq(0.233);
Acer negundo	NA	Vran	0.876	0.989	0.98	NA	NA	Tminw(0.142);Twetq(0.2);Tcoldq(0.769);Pa(0.301);Pvar(0.091);Pwarmq(0.086);Pcoldq(0.058);Mcoldq(0.124);
Asclepsias syriaca	EU	Vran	0.903	0.975	0.892	0.993	0.802	Tvar(0.55);Twarmq(0.21);Pa(0.078);Pwarmq(0.232);Ma(0.096);Mminw(0.384);Mdryq(0.253);Mwarmq(0.305);
Aster noviibelgi	EU	Vran	0.89	0.942	0.985	0.981	0.972	Tvar(0.328);Pwetw(0.112);Pdryw(0.287);Pvar(0.111);Pwetq(0.323);Pcoldq(0.131);Ma(0.068);Mwarmq(0.227);
Bidens frondosa	EU	Vran	0.908	0.94	0.952	0.978	0.955	isoT(0.56);Tmaxw(0.181);Pwetw(0.131);Pwarmq(0.066);Ma(0.101);Mvar(0.142);Mwarmq(0.143);Mcoldq(0.711);
Epilobium ciliatum	EU	Vran	0.782	0.954	0.917	0.974	0.991	isoT(0.283);Pwetw(0.288);Pvar(0.051);Pdryq(0.228);Mmaxw(0.13);Mvar(0.214);Mwetq(0.328);Mwarmq(0.197);
Helianthus tuberosus[1]	EU	Vran	0.891	0.992	0.934	0.972	0.975	isoT(0.23);Tvar(0.269);Tminw(0.684);Tarange(0.262);Pwetw(0.168);Pcoldq(0.043);Mvar(0.205);Mcoldq(0.285);
Rudbeckia laciniata	EU	Vran	0.856	0.986	0.976	0.945	0.988	Tdrange(0.019);Tvar(0.438);Tmaxw(0.119);Pa(0.189);Pdryw(0.167);Pvar(0.067);Ma(0.02);Mwarmq(0.137);

Table 2: Description of Variables

Ref. number	Abbreviation	Description
1	Tmean	Annual mean temperature (°C)
2	Tdrange	Mean diurnal temperature range (mean(period max-min)) (°C)
3	IsoT	Isothermality (Bio02 ÷ Bio07)
4	Tvar	Temperature seasonality (C of V)
5	Tmaxw	Max temperature of warmest week (°C)
6	Tcoldw	Min temperature of coldest week (°C)
7	Tarange	Temperature annual range (Bio05-Bio06) (°C)
8	Twetq	Mean temperature of wettest quarter (°C)
9	Tdryq	Mean temperature of driest quarter (°C)
10	Twarmq	Mean temperature of warmest quarter (°C)
11	Tcoldq	Mean temperature of coldest quarter (°C)
12	Pa	Annual precipitation (mm)
13	Pwetw	Precipitation of wettest week (mm)
14	Pdryw	Precipitation of driest week (mm)
15	Pvar	Precipitation seasonality (C of V)
16	Pwetq	Precipitation of wettest quarter (mm)
17	Pdryq	Precipitation of driest quarter (mm)
18	Pwarmq	Precipitation of warmest quarter (mm)
19	Pcoldq	Precipitation of coldest quarter (mm)
20	Ma	Annual mean moisture index
21	Mwetw	Highest weekly moisture index
22	Mdryw	Lowest weekly moisture index
23	Mvar	Moisture index seasonality (C of V)
24	Mwetq	Mean moisture index of wettest quarter
25	Mdryq	Mean moisture index of driest quarter
26	Mwarmq	Mean moisture index of warmest quarter
27	Mcoldq	Mean moisture index of coldest quarter
28-35	PC	Principal components calibrated on the 27 variables

Table 3: Species traits and p-value in the MANOVA-test with variables contribution of the best transferable SDM when including only non-shifting species (*P*).

Trait	Levels	Categories	P
Species origin	EU, NA	Distribution	0.147
Time since 1st introduction (in Holarctic)	Continuous	Distribution	0.321
Native niche breadth	Continuous	Distribution	0.757
Main Biotop	Forest/shrubland, ruderal, river/lakeshore, fertile meadows	Distribution	0.736
Perennial	Yes, No	Reproduction	0.435
Vegetative reproduction	Yes, No	Reproduction	0.669
Outcrossing	Yes, Rare, No	Reproduction	0.943
Selfing	Yes, Rare, No	Reproduction	0.989
C Grime	Yes, No	CSR	0.755
S Grime	Yes, No	CSR	0.969
R Grime	Yes, No	CSR	0.869
Growth Habit	Tree, Shrub, Forb, Tree	Competitivity	0.994
Height	Continuous	Competitivity	0.956
Invasiveness	Invasive, Weed	Competitivity	0.481
Ploidy	Diploid, Triploid or more (up to octoploid)	Competitivity	0.206

Table 4: Traits explaining SDM transferability expressed by Boyce index (B) and Sensitivity (Se) for Vsoa and Vpc2 in Holarctic and Australian invaded range. Niche shifting species were not included in the analysis.

	Holarctic	**Australia**
V_{soa} Se	native niche breadth (+)	Species origin (-), Niche breadth (+)
V_{soa} B	S Grime (+)	Species Origin (-)
V_{PC2} Se	Species origin (+), height (+), native niche breadth (+)	Repro.veg (-)
V_{PC2} B	Ploidy (-),native niche breadth (+)	Species Origin (-)

REFERENCES

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

Araújo, M.B., Ferri-Yáñez, F., Bozinovic, F., Marquet, P.A., Valladares, F. & Chown, S.L. (2013) Heat freezes niche evolution. *Ecology Letters*, **16,** 1206-1219.

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

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

Araújo, M.B. & Rahbek, C. (2006) How Does Climate Change Affect Biodiversity? *Science,* **313,** 1396-1397.

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

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

Austin, M.P. & Van Niel, K.P. (2011) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography,* **38,** 1-8.

Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012) Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution,* **3,** 327-338.

Barton, K. (2012) MuMIn: Multi-model inference.

Bosshard, T., Kotlarski, S., Ewen, T. & Schär C. (2011) Spectral representation of the annual cycle in the climate change signal. *Hydrological & Earth System Sciences Discussions,* **8,** 1161-1192.

Braunisch, V., Coppes, J., Arlettaz, R., Suchant, R., Schmid, H. & Bollmann, K. (2013) Selecting from correlated climate variables: a major source of uncertainty for predicting species distributions under climate change. *Ecography,* **36,** 971-983.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters,* **10,** 701-709.

Buckley, L.B., Urban, M.C., Angilletta, M.J., Crozier, L.G., Rissler, L.J. & Sears, M.W. (2010) Can mechanism inform species' distribution models? *Ecology Letters,* **13,** 1041-1054.

Caplat, P., Cheptou, P.O., Diez, J., Guisan, A., Larson, B., Macdougall, A., Peltzer, D., Richardson, D., Shea, K. & van Kleunen, M. (2013) Movement, impacts and management of plant distributions in response to climate change: insights from invasions. *Oikos,* **122,** 1265-1274.

Colautti, R., Grigorovich, I. & MacIsaac, H. (2006) Propagule Pressure: A Null Model for Biological Invasions. *Biological Invasions,* **8,** 1023-1037.

Cutler, D.R., Edwards, T.C., Beard, K.H., Cutler, A. & Hess, K.T. (2007) Random forests for classification in ecology. *Ecology,* **88,** 2783-2792.

Daehler, C.C. (2003) Performance comparisons of co-occurring native and alien invasive plants: Implications for conservation and restoration. *Annual Review of Ecology Evolution and Systematics,* **34,** 183-211.

Dormann, C.F. (2007) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography,* **16,** 129-138.

Dormann, C.F., Purschke, O., Marquez, J.R.G., Lautenbach, S. & Schroder, B. (2008) Components of uncertainty in species distribution analysis: a case study of the great grey shrike. *Ecology,* **89,** 3371-3386.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution***, 1,** 330-342 .

Engler, R., Randin, C.F., Thuiller, W., Dullinger, S., Zimmermann, N.E., Araujo, M.B., Pearman, P.B., Le Lay, G., Piedallu, C., Albert, C.H., Choler, P., Coldea, G., De Lamo, X., Dirnbock, T., Gegout, J.C.,

Gomez-Garcia, D., Grytnes, J.A., Heegaard, E., Hoistad, F., Nogues-Bravo, D., Normand, S., Puscas, M., Sebastia, M.T., Stanisci, A., Theurillat, J.P., Trivedi, M.R., Vittoz, P. & Guisan, A. (2011) 21st century climate change threatens mountain flora unequally across Europe. *Global Change Biology,* **17,** 2330-2341.

Fitzpatrick, M.C. & Hargrove, W.W. (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation,* **18,** 2255-2261.

Friedman, J.H., Hastie, T.J. & Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting. *Annals of Statistics,* **28,** 337-374.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology,* **98,** 790-799.

Giovanelli, J.G.R., de Siqueira, M.F., Haddad, C.F.B. & Alexandrino, J. (2010) Modeling a spatially restricted distribution in the Neotropics: How the size of calibration area affects the performance of five presence-only methods. *Ecological Modelling,* **221,** 215-224.

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

Green, R.H. (1979) *Sampling Design and Statistical Methods for Environmental Biologists.* John Wiley, New York.

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

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters,* **16,** 1424-1435.

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

Harrell, F.E., Lee, K.L., Califf, R.M., Pryor, D.B. & Rosati, R.A. (1984) Regression modelling strategies for improved prognostic prediction. *Statistics in Medicine,* **3,** 143-152.

Hellmann, J.J., Byers, J.E., Bierwagen, B.G. & Dukes, J.S. (2008) Five Potential Consequences of Climate Change for Invasive Species. *Conservation Biology,* **22,** 534-543.

Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2010) dismo: Species distribution modeling. pp. R package version 0.7-17. Available at: http://CRAN.R-project.org/package=dismo. Last accessed 8 August 2012.

Hill, M.P., Hoffmann, A.A., Macfadyen, S., Umina, P.A. & Elith, J. (2012) Understanding niche shifts: using current and historical data to model the invasive redlegged earth mite, *Halotydeus destructor*. *Diversity and Distributions,* **18,** 191-203.

Hulme, P.E. & Barrett, S.C.H. (2013) Integrating trait- and niche-based approaches to assess contemporary evolution in alien plant species. *Journal of Ecology,* **101,** 68-77.

Hutchinson, G.E. (1957) Population Studies - Animal Ecology and Demography - Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology,* **22,** 415-427.

Jeschke, J.M. & Strayer, D.L. (2008) Usefulness of bioclimatic models for studying climate change and invasive species. *Year in Ecology and Conservation Biology 2008*, pp. 1-24.

Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters,* **12,** 334-350.

Koop, A., Fowler, L., Newton, L. & Caton, B. (2012) Development and validation of a weed screening tool for the United States. *Biological Invasions,* **14,** 273-294.

Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J. & Scott, J.K. (2011) CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution,* **3,** 53-64.

Kueffer, C., Pyšek, P. & Richardson, D.M. (2013) Integrative invasion science: model systems, multi-site studies, focused meta-analysis and invasion syndromes. *New Phytologist,* **200,** 615-633.

Kuhn, I., Durka, W. & Klotz, S. (2004) BiolFlor - a new plant-trait database as a tool for plant invasion ecology. *Diversity and Distributions,* **10,** 363-365.

Lenz, A., Hoch, G., Vitasse, Y. & Körner, C. (In Press) European deciduous trees exhibit similar safety margins against damage by spring freeze events along elevational gradients. *New Phytologist***, 200,** 1166-1175.

Mac Nally, R. (2002) Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables. *Biodiversity and Conservation,* **11,** 1397-1401.

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

Mainka, S.A. & Howard, G.W. (2010) Climate change and invasive species: double jeopardy. *Integrative Zoology,* **5,** 102-111.

Maiorano, L., Cheddadi, R., Zimmermann, N.E., Pellissier, L., Petitpierre, B., Pottier, J., Laborde, H., Hurdu, B.I., Pearman, P.B., Psomas, A., Singarayer, J.S., Broennimann, O., Vittoz, P., Dubuis, A., Edwards, M.E., Binney, H.A. & Guisan, A. (2012) Building the niche through time: using 13,000 years of data to predict the effects of climate change on three tree species in Europe. *Global Ecology and Biogeography,* **22,** 302-317.

Mandle, L., Warren, D.L., Hoffmann, M.H., Peterson, A.T., Schmitt, J. & von Wettberg, E.J. (2010) Conclusions about Niche Expansion in Introduced Impatiens walleriana Populations Depend on Method of Analysis. *Plos One,* **5**.

Martinez-Meyer, E. & 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.

McCullagh, P. & Nelder, J.A. (1983) *Generalized Linear Models: First Edition.* Chapman and Hall, London, New-York.

McPherson, J.M. & Jetz, W. (2007) Effects of species' ecology on the accuracy of distribution models. *Ecography,* **30,** 135-151.

Mellin, C., Bradshaw, C.J.A., Meekan, M.G. & Caley, M.J. (2010) Environmental and spatial predictors of species richness and abundance in coral reef fishes. *Global Ecology and Biogeography,* **19,** 212-222.

Meynard, C.N., Migeon, A. & Navajas, M. (2013) Uncertainties in Predicting Species Distributions under Climate Change: A Case Study Using *Tetranychus evansi* (Acari: Tetranychidae), a Widespread Agricultural Pest. *Plos One,* **8,** e66445.

Moles, A.T., Gruber, M.A.M. & Bonser, S.P. (2008) A new framework for predicting invasive plant species. *Journal of Ecology,* **96,** 13-17.

Morin, X. & Thuiller, W. (2009) Comparing niche- and process-based models to reduce prediction uncertainty in species range shifts under climate change. *Ecology,* **90,** 1301-1313.

Normand, S., Treier, U.A., Randin, C., Vittoz, P., Guisan, A. & Svenning, J.C. (2009) Importance of abiotic stress as a range-limit determinant for European plants: insights from species responses to climatic gradients. *Global Ecology and Biogeography,* **18,** 437-449.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008*a*) Niche dynamics in space and time. *Trends in Ecology & Evolution,* **23,** 149-158.

Pearman, P.B., Randin, C.F., Broennimann, O., Vittoz, P., van der Knaap, W.O., Engler, R., Le Lay, G., Zimmermann, N.E. & Guisan, A. (2008*b*) Prediction of plant species distributions across six millennia. *Ecology Letters,* **11,** 357-369.

Pellissier, L., Bråthen, K.A., Vittoz, P., Yoccoz, N.G., Dubuis, A., Meier, E.S., Zimmermann, N.E., Randin, C.F., Thuiller, W., Garraud, L., Van Es, J. & Guisan, A. (2013a) Thermal niches are more conserved at cold than warm limits in arctic-alpine plant species. *Global Ecology and Biogeography,* **22,** 933-941.

Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology,* **78,** 419-433.

Peterson, A.T., Papes, M. & Eaton, M. (2007) Transferability and model evaluation in ecological niche modeling: a comparison of GARP and Maxent. *Ecography,* **30,** 550-560.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography,* **38,** 817-827.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders. *Science,* **335,** 1344-1348.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* **190,** 231-259.

Phillips, S.J. & Elith, J. (2010) POC plots: calibrating species distribution models with presence-only data. *Ecology,* **91,** 2476-2484.

Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C.F., Vittoz, P. & Guisan, A. (2013) The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Global Ecology and Biogeography,* **22,** 52-63.

Pyke, C.R., Thomas, R., Porter, R.D., Hellmann, J.J., Dukes, J.S., Lodge, D.M. & Chavarria, G. (2008) Current Practices and Future Opportunities for Policy on Climate Change and Invasive Species. *Conservation Biology,* **22,** 585-592.

R (2005) *R : A language and environment for statistical computing. Version 2.5.1.* Development Core Team. R Foundation for Statistical Computing, Vienna, Austria.

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

Randin, C.F., Paulsen, J., Vitasse, Y., Kollas, C., Wohlgemuth, T., Zimmermann, N.E. & Körner, C. (2013) Do the elevational limits of deciduous tree species match their thermal latitudinal limits? *Global Ecology and Biogeography,* **22,** 913-923.

Rodder, D. & Lotters, S. (2010) Explanative power of variables used in species distribution modelling: an issue of general model transferability or niche shift in the invasive Greenhouse frog (*Eleutherodactylus planirostris*). *Naturwissenschaften,* **97,** 781-796.

Rodder, D., Schmidtlein, S., Veith, M. & Lotters, S. (2009) Alien Invasive Slider Turtle in Unpredicted Habitat: A Matter of Niche Shift or of Predictors Studied? *Plos One,* **4**.

Roura-Pascual, N., Brotons, L., Peterson, A.T. & Thuiller, W. (2009) Consensual predictions of potential distributional areas for invasive species: a case study of Argentine ants in the Iberian Peninsula. *Biological Invasions,* **11,** 1017-1031.

Savolainen, O., Bokma, F., GarcÄ±Ì a-Gil, R., Komulainen, P.i. & Repo, T. (2004) Genetic variation in cessation of growth and frost hardiness and consequences for adaptation of *Pinus sylvestris* to climatic changes. *Forest Ecology and Management,* **197,** 79-89.

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

Soberón, J. & Peterson, A.T. (2011) Ecological niche shifts and environmental space anisotropy: a cautionary note. *Revista Mexicana de Biodiversidad,* **82,** 1348-1355.

Thompson, K. & Davis, M.A. (2011) Why research on traits of invasive plants tells us very little. *Trends in ecology & evolution (Personal edition),* **26,** 155-156.

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

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (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.

Vale, C.G., Tarroso, P. & Brito, J.C. (2013) Predicting species distribution at range margins: testing the effects of study area extent, resolution and threshold selection in the Sahara–Sahel transition zone. *Diversity and Distributions***, 20,** 20-33.

Van Kleunen, M., Weber, E. & Fischer, M. (2010) A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters,* **13,** 235-245.

Venette, R.C., Kriticos, D.J., Magarey, R.D., Koch, F.H., Baker, R.H., Worner, S.P., Raboteaux, N.N.G., McKenney, D.W., Dobesberger, E.J. & Yemshanov, D. (2010) Pest risk maps for invasive alien species: a roadmap for improvement. *BioScience,* **60,** 349-362.

Verbruggen, H., Tyberghein, L., Belton, G.S., Mineur, F., Jueterbock, A., Hoarau, G., Gurgel, C.F.D. & De Clerck, O. (2013) Improving Transferability of Introduced Species Distribution Models: New Tools to Forecast the Spread of a Highly Invasive Seaweed. *Plos One,* **8**.

Ward, D. (2007) Modelling the potential geographic distribution of invasive ant species in New Zealand. *Biological Invasions,* **9,** 723-735.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. *Evolution,* **62,** 2868-2883.

Warren, D.L. & Seifert, S.N. (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications,* **21,** 335-342.

Warren, M., Robertson, M.P. & Greeff, J.M. (2010) A comparative approach to understanding factors limiting abundance patterns and distributions in a fig tree-fig wasp mutualism. *Ecography,* **33,** 148-158.

Whitney, K.D. & Gabler, C.A. (2008) Rapid evolution in introduced species, 'invasive traits' and recipient communities: challenges for predicting invasive potential. *Diversity and Distributions,* **14,** 569-580.

Williams, K.J., Belbin, L., Austin, M.P., Stein, J.L. & Ferrier, S. (2012) Which environmental variables should I use in my biodiversity model? *International Journal of Geographical Information Science*, **26**, 2009-2047.

Williamson, M. (1999) Invasions. *Ecography,* **22,** 5-12.

Woodward, F.I. (1987) *Climate and plant distribution.* Cambridge University Press, Cambridge, UK.

Zanini, F., Pellet, J. & Schmidt, B.R. (2009) The transferability of distribution models across regions: an amphibian case study. *Diversity and Distributions,* **15,** 469-480.

Zhang, Q. & Zhang, X. (2011) Impacts of predictor variables and species models on simulating *Tamarix ramosissima* distribution in Tarim Basin, northwestern China. *Journal of Plant Ecology,* **5,** 337-345.

Zweig, M.H. & Campbell, G. (1993) Receiver-Operating Characteristic (Roc) Plots - a Fundamental Evaluation Tool in Clinical Medicine. *Clinical Chemistry,* **39,** 561-577.

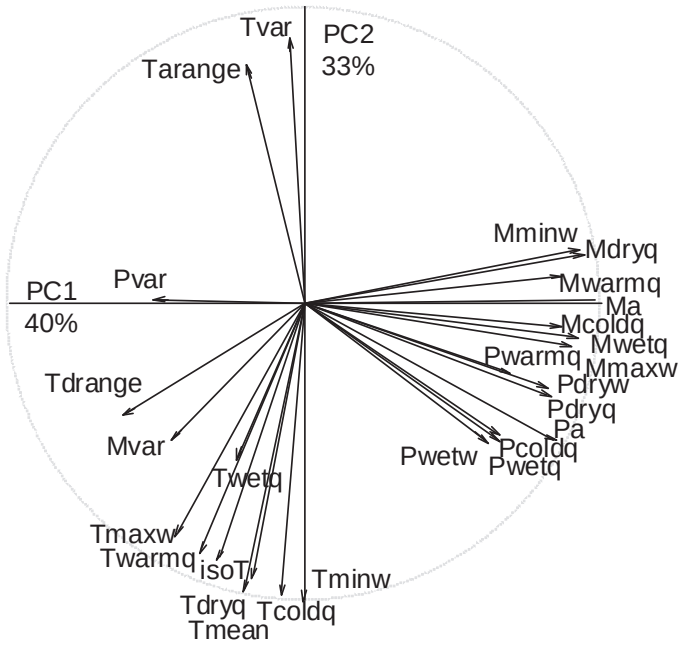


Figure S1: correlation between the two first components of the PCA and the 27 variables. Proportion of explained variances is indicated for both axes

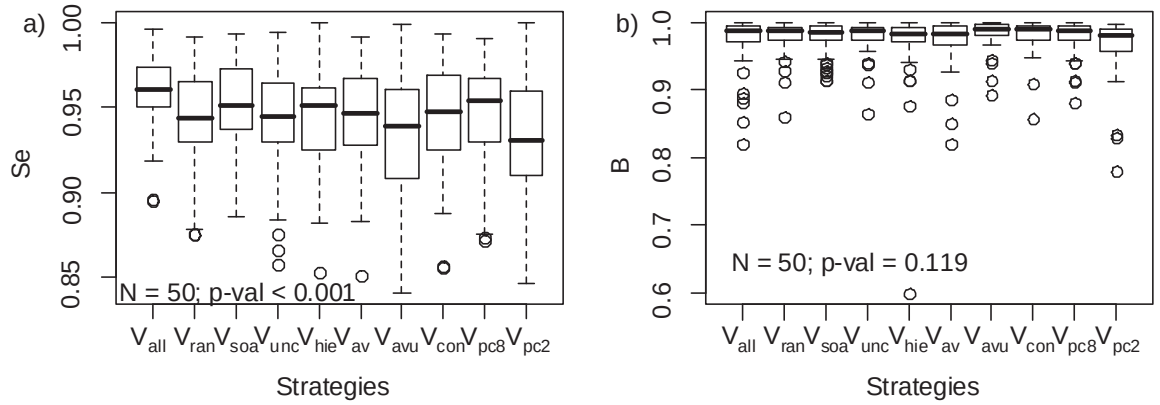


Figure S2: Sensitivity (Se) and Boyce index (B) calculated on the pseudo-independent dataset in the native range

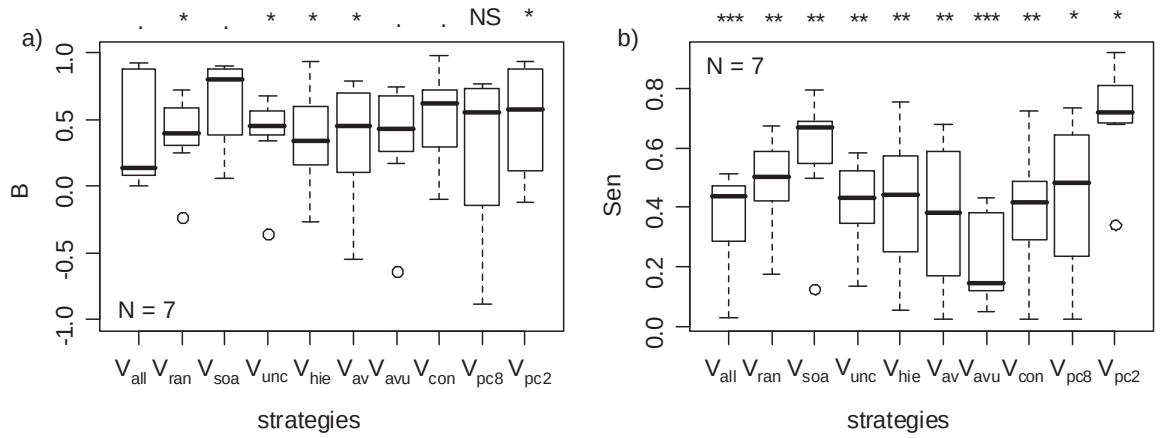


Figure S3: Evaluation for Holarctic shifting species in the invaded range following different variables strategies. Models are evaluated with Boyce index (a) and sensitivity (b) when projected onto their reciprocal Holarctic invaded range (EU or NA). For each strategy, we tested if the performance were lower than for non shifting species and indicated the p-value or a t-test**.**

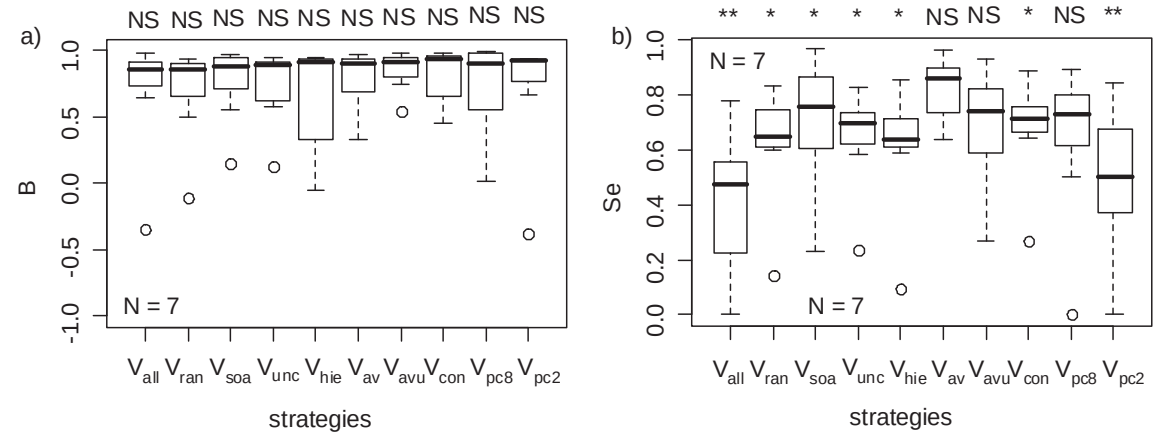


Figure S4: Evaluation for Australian shifting species in the invaded range following different variables strategies. Models are evaluated with Boyce index (a) and sensitivity (b) when projected onto their reciprocal Holarctic invaded range (EU or NA). For each strategy, we tested if the performance were lower than for non shifting species and indicated the p-value or a t-test.

See attached file

https://docs.google.com/file/d/0B5ZZgCyoOEdnaHJoYmdyem5DS00/edit?usp=sharing

Figure S5 : projections of the best model

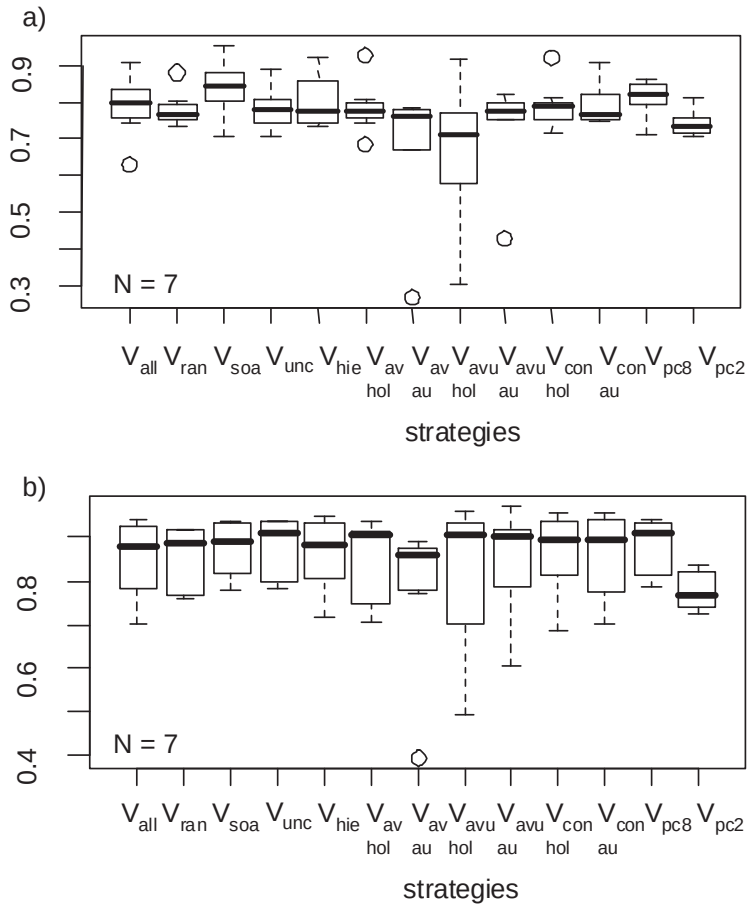


Figure S6: Pearson's correlation between best models projections under future climate in EU, NA and AU and projections provided by variables selection strategies for niche shifting species in Holarctic (a) and Australia (b).

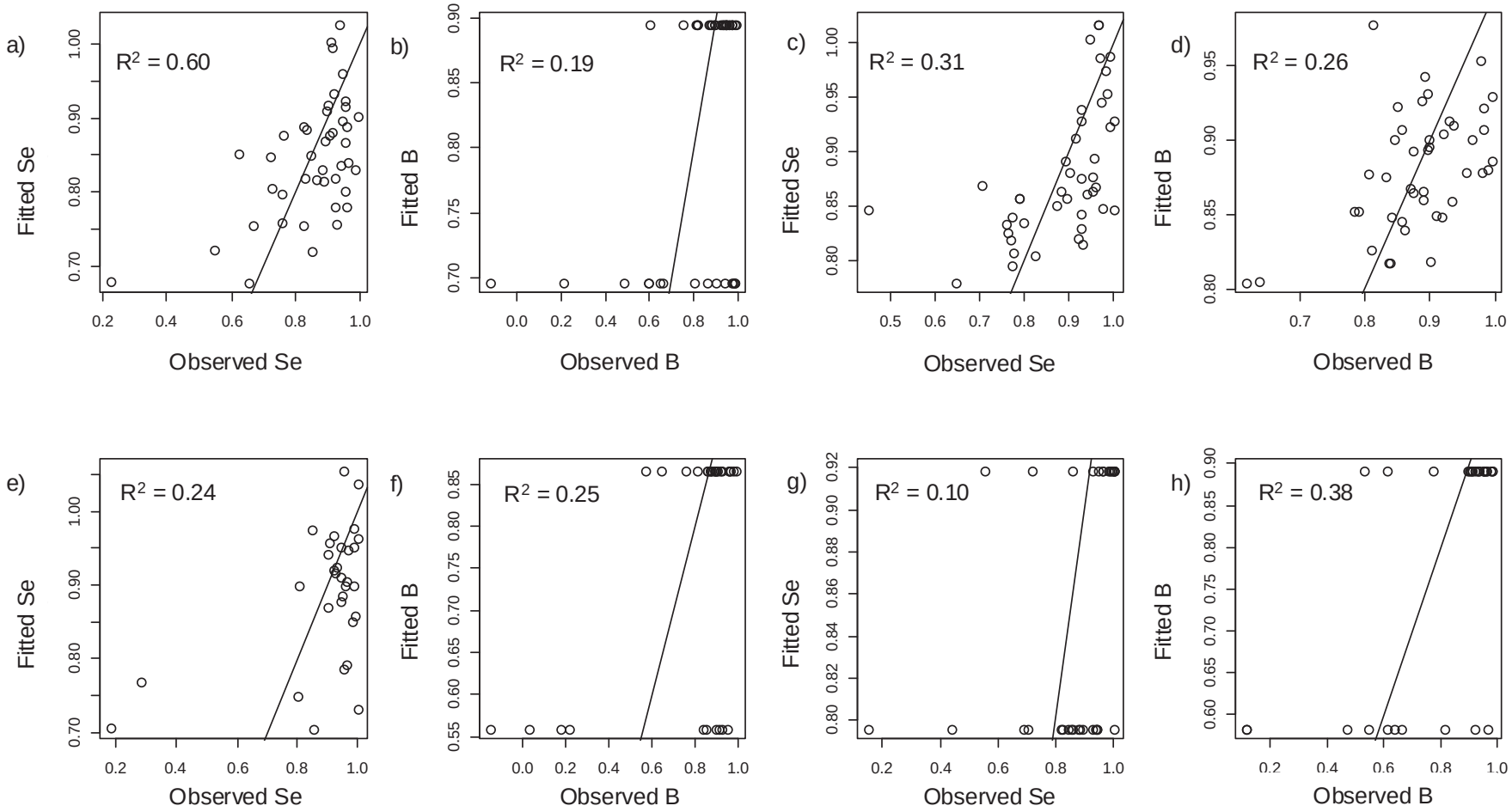


Figure S7: Fitted vs observed value when transferability is correlated to species traits. Transferability was assessed with Sensitivity (Se) and Boyce (B) in the Holarctic (a, b, c, d) and Australian (e, f, g, h) invaded ranges for non-shifting species. For each model, evaluation of fit is provided by the proportion of explained deviance (R^2).

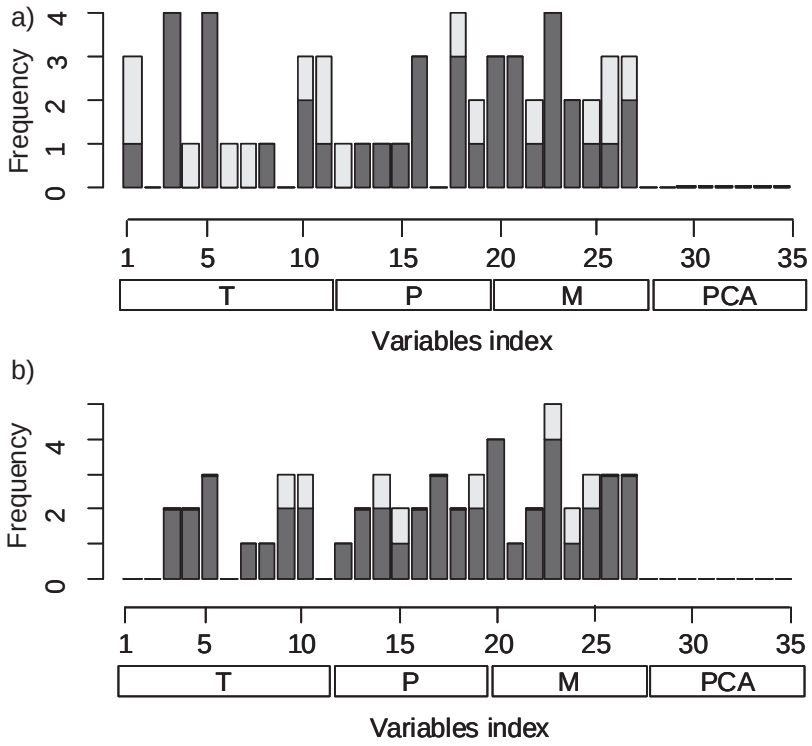


Figure S8: Number of times variables are included in the best model for niche shifting species in Holarctic (a) and Australia (b). Black and grey represent EU and NA species respectively. The index is the same as Table 1. T, P, M and PCA represent temperature, precipitation, moisture and principal component variables

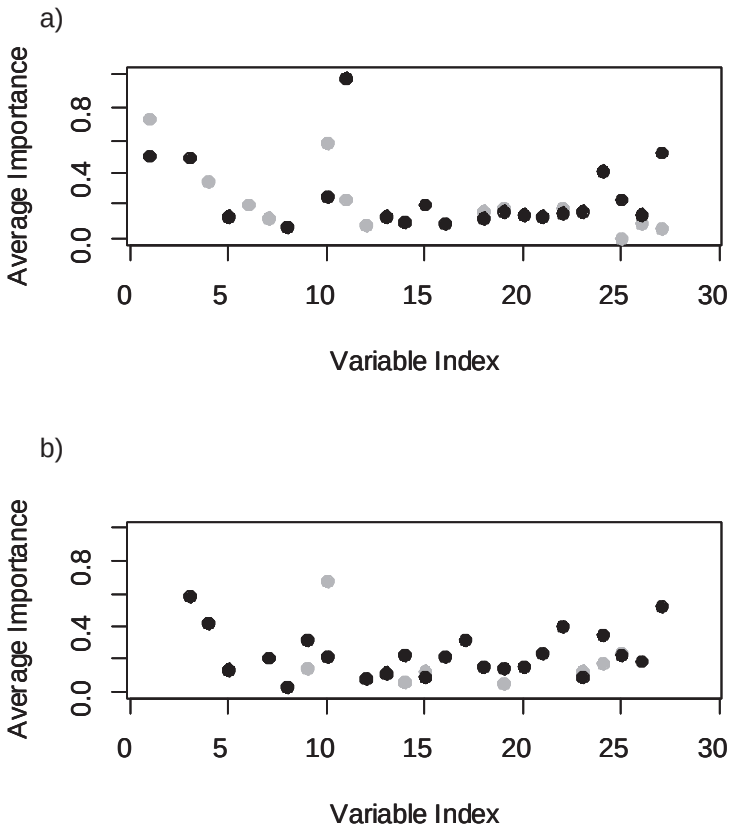


Figure S9: Average variables importance once they are included in the best model for niche shifting species in Holarctic (a) and Australia (b). Black and grey represent EU and NA species respectively. T, P, M and PC represent temperature, precipitation, moisture and principal component variables.

See attached file

https://docs.google.com/file/d/0B5ZZgCyoOEdnQ0hxMk9oODU0ZDA/edit?usp=sharing

Figure S10: Response curve of the ENSEMBLE best model for each species

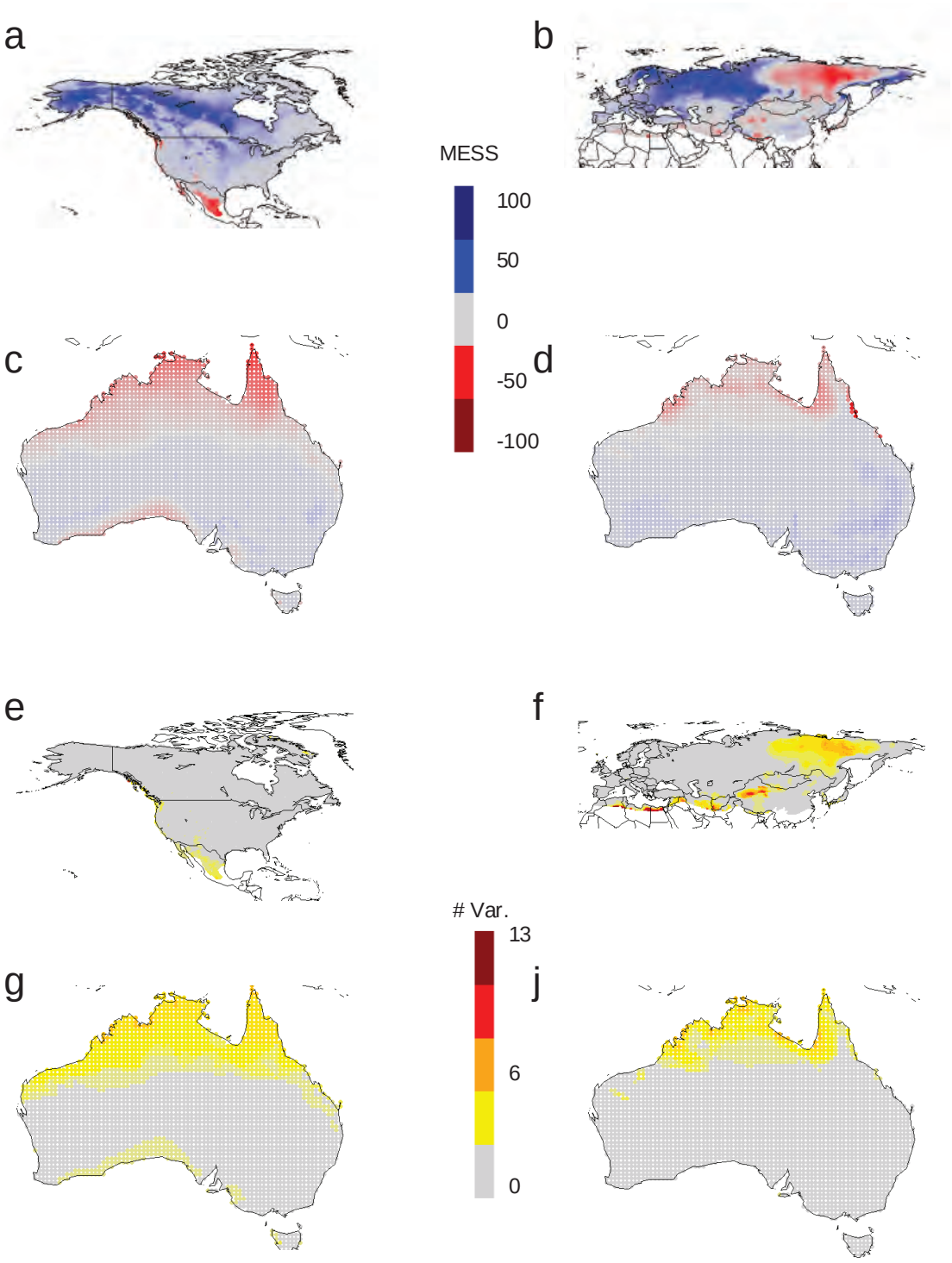


Figure S11: MESS values (a-d) and number of variables showing non analog climate with the native range (e-h) with the native Eurasian range (a,c,e,g) and native North American range (b,d,f,h)

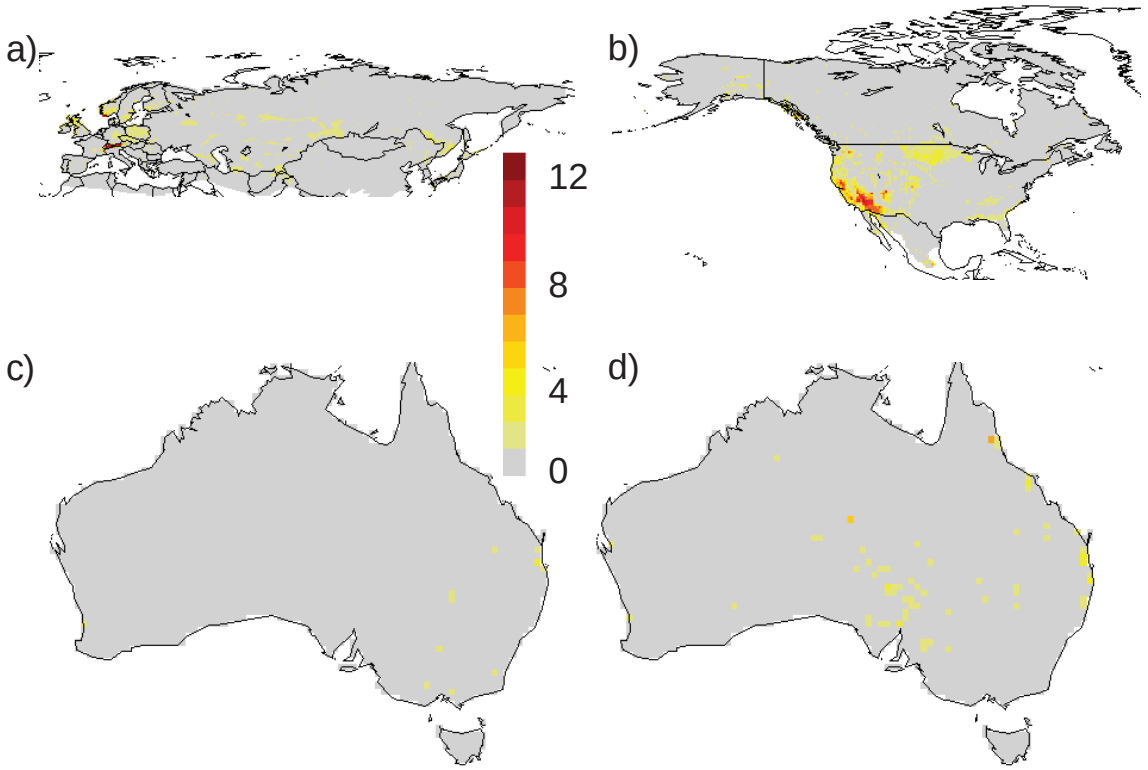


Figure S12: Number of species showing expansion (i.e. predicted as absence by the best model but where presences occur) in their invaded range. A) and c) represent invaded range of Eurasian species and b) and d) represent invaded ranges of North American species.

Table S1: TSS in the native range for each species and strategy. Average TSS among 1000 replicates is provided for V_{ran} and V_{unc}

Species	Nat	Inv	V_{all}	V_{ran}	V_{soa}	V_{unc}	V_{hie}	V_{av}	V_{avu}	V_{con}	V_{pc8}	V_{pc2}
Alliaria petiolata	eu	na	0.899	0.858	0.873	0.864	0.858	0.854	0.864	0.862	0.868	0.789
Amaranthus retroflexus	na	eu	0.922	0.887	0.909	0.882	0.899	0.89	0.827	0.891	0.893	0.72
Ambrosia artemisiifolia	na	eu	0.9	0.87	0.868	0.868	0.877	0.858	0.831	0.88	0.871	0.727
Amorpha fruticosa	na	eu	0.78	0.754	0.743	0.756	0.775	0.747	0.722	0.753	0.781	0.731
Anagallis arvensis	eu	na	0.932	0.912	0.921	0.914	0.92	0.913	0.914	0.916	0.917	0.849
Anthoxanthum odoratum	eu	na	0.918	0.886	0.892	0.888	0.891	0.893	0.888	0.893	0.888	0.8
Arabidopsis thaliana	eu	na	0.905	0.867	0.881	0.866	0.868	0.882	0.869	0.879	0.881	0.723
Bromus sterilis	eu	na	0.951	0.931	0.943	0.932	0.931	0.935	0.931	0.934	0.94	0.874
Bromus tectorum	eu	na	0.925	0.886	0.905	0.887	0.893	0.898	0.893	0.893	0.897	0.827
Carduus nutans	eu	na	0.864	0.836	0.851	0.838	0.835	0.84	0.836	0.841	0.844	0.702
Centaurea stoebe	eu	na	0.919	0.909	0.921	0.909	0.91	0.917	0.919	0.915	0.918	0.888
Cirsium vulgare	eu	na	0.938	0.903	0.917	0.904	0.913	0.917	0.903	0.913	0.915	0.763
Conyza canadensis	na	eu	0.949	0.927	0.922	0.926	0.922	0.932	0.912	0.928	0.937	0.899
Cytisus scoparius	eu	na	0.963	0.938	0.955	0.94	0.949	0.945	0.95	0.951	0.942	0.867
Dactylis glomerata	eu	na	0.891	0.833	0.848	0.834	0.84	0.851	0.832	0.834	0.832	0.67
Echinocystis lobata	na	eu	0.915	0.877	0.883	0.865	0.883	0.885	0.812	0.885	0.865	0.727
Erigeron annuus	na	eu	0.961	0.944	0.954	0.94	0.949	0.946	0.902	0.957	0.929	0.828
Erodium cicutarium	eu	na	0.916	0.872	0.89	0.877	0.888	0.871	0.879	0.885	0.89	0.765
Euphorbia esula	eu	na	0.884	0.845	0.86	0.848	0.845	0.861	0.85	0.852	0.853	0.762
Holcus lanatus	eu	na	0.96	0.934	0.949	0.935	0.934	0.933	0.937	0.931	0.947	0.861
Hypochaeris radicata	eu	na	0.974	0.953	0.965	0.956	0.959	0.961	0.961	0.956	0.961	0.882
Juncus tenuis	na	eu	0.926	0.871	0.888	0.863	0.888	0.874	0.793	0.898	0.852	0.689
Linaria vulgaris	eu	na	0.906	0.853	0.875	0.856	0.856	0.873	0.861	0.86	0.862	0.728
Lythrum salicaria	eu	na	0.935	0.887	0.911	0.893	0.904	0.897	0.894	0.903	0.887	0.707
Medicago lupulina	eu	na	0.86	0.755	0.78	0.749	0.77	0.78	0.73	0.763	0.755	0.629
Melilotus albus	eu	na	0.878	0.781	0.814	0.777	0.806	0.817	0.765	0.813	0.777	0.63
Phytolacca americana	na	eu	0.857	0.855	0.865	0.852	0.857	0.862	0.798	0.858	0.846	0.825
Plantago lanceolata	eu	na	0.9	0.854	0.87	0.859	0.872	0.866	0.86	0.87	0.857	0.715
Plantago major	eu	na	0.841	0.781	0.782	0.784	0.791	0.792	0.769	0.777	0.796	0.658
Poa annua	eu	na	0.87	0.811	0.826	0.811	0.815	0.814	0.798	0.817	0.822	0.652
Potentilla recta	eu	na	0.913	0.869	0.879	0.871	0.867	0.88	0.87	0.875	0.873	0.809
Prunus serotina	na	eu	0.947	0.925	0.943	0.926	0.931	0.931	0.886	0.928	0.93	0.849
Rhus typhina	na	eu	0.964	0.951	0.963	0.951	0.959	0.957	0.92	0.951	0.946	0.888
Robinia pseudoacacia	na	eu	0.828	0.8	0.799	0.796	0.797	0.811	0.731	0.808	0.793	0.735
Rumex acetosella	eu	na	0.901	0.839	0.875	0.838	0.85	0.86	0.838	0.841	0.858	0.677
Solidago canadensis	na	eu	0.87	0.828	0.848	0.822	0.845	0.839	0.757	0.844	0.831	0.674
Solidago gigantea	na	eu	0.928	0.877	0.878	0.876	0.89	0.89	0.852	0.886	0.874	0.792
Sonchus oleraceus	eu	na	0.928	0.899	0.917	0.904	0.908	0.91	0.902	0.902	0.909	0.795
Trifolium arvensis	eu	na	0.919	0.888	0.898	0.893	0.884	0.889	0.892	0.89	0.901	0.751
Trifolium dubium	eu	na	0.961	0.936	0.951	0.933	0.932	0.948	0.937	0.933	0.948	0.875
Trifolium repens	eu	na	0.872	0.806	0.845	0.805	0.814	0.829	0.798	0.812	0.841	0.659
Verbascum thapsus	eu	na	0.925	0.871	0.899	0.869	0.871	0.887	0.868	0.874	0.882	0.737
Vicia sativa	eu	na	0.905	0.856	0.878	0.863	0.872	0.865	0.874	0.873	0.863	0.749
Acer negundo	na	eu	0.883	0.841	0.859	0.839	0.855	0.849	0.767	0.851	0.844	0.737
Asclepsias syriaca	na	eu	0.919	0.91	0.911	0.907	0.915	0.912	0.888	0.915	0.902	0.801
Aster noviibelgi	na	eu	0.939	0.918	0.94	0.918	0.915	0.913	0.926	0.922	0.918	0.896
Bidens frondosa	na	eu	0.914	0.858	0.845	0.855	0.852	0.853	0.838	0.851	0.846	0.792
Epilobium ciliatum	na	eu	0.697	0.653	0.659	0.644	0.655	0.675	0.609	0.663	0.622	0.547
Helianthus tuberosus	na	eu	0.924	0.902	0.91	0.897	0.908	0.89	0.828	0.901	0.881	0.752
Rudbeckia laciniata	na	eu	0.877	0.836	0.835	0.833	0.85	0.848	0.793	0.835	0.834	0.723

Table S2: Sensitivity in the invaded Holarctic range for each species and strategy. Average Sensitivity among 1000 replicates is provided for V_{ran} and V_{unc}.

Species	Nat	V_{all}	V_{ran}	V_{soa}	V_{unc}	V_{hie}	V_{av}	V_{avu}	V_{con}	V_{pc8}	V_{pc2}
Alliaria petiolata	eu	0.889	0.919	0.889	0.928	0.96	0.874	0.886	0.932	0.895	0.975
Amaranthus retroflexus	na	0.545	0.503	0.724	0.412	0.633	0.567	0.061	0.365	0.576	0.89
Ambrosia artemisiifolia	na	0.729	0.612	0.936	0.581	0.72	0.602	0.062	0.611	0.495	0.913
Amorpha fruticosa1	na	0.452	0.445	0.677	0.342	0.387	0.548	0.097	0.355	0.161	0.677
Anagallis arvensis	eu	0.948	0.877	0.907	0.881	0.781	0.79	0.656	0.577	0.988	0.991
Anthoxanthum odoratum	eu	0.794	0.922	0.962	0.888	0.987	0.955	0.87	0.972	0.947	0.925
Arabidopsis Thaliana	eu	0.925	0.888	0.992	0.905	0.951	0.992	0.971	0.965	0.909	0.952
Bromus sterilis1	eu	0.157	0.523	0.798	0.468	0.753	0.382	0.337	0.416	0.674	0.787
Bromus tectorum	eu	0.693	0.616	0.618	0.601	0.558	0.528	0.526	0.529	0.668	0.795
Carduus nutans	eu	0.646	0.776	0.862	0.809	0.846	0.923	0.692	0.846	0.785	0.769
Centaurea stoebe1	eu	0.027	0.173	0.123	0.135	0.053	0.023	0.049	0.024	0.307	0.339
Cirsium vulgare2	eu	0.664	0.774	0.882	0.773	0.775	0.885	0.868	0.898	0.639	0.918
Conyza canadensis	na	0.851	0.82	0.909	0.771	0.946	0.914	0.371	0.849	0.891	0.946
Cytisus scoparius1	eu	0.419	0.501	0.5	0.433	0.113	0.145	0.145	0.226	0.484	0.694
Dactylis glomerata	eu	0.929	0.897	0.95	0.913	0.88	0.938	0.917	0.906	0.951	0.899
Echinocystis lobata	na	0.524	0.558	0.752	0.521	0.402	0.677	0.162	0.49	0.618	0.955
Erigeron annuus	na	0.824	0.565	0.85	0.471	0.68	0.857	0.252	0.699	0.189	0.926
Erodium cicutarium	eu	0.567	0.653	0.898	0.554	0.503	0.812	0.482	0.554	0.794	0.927
Euphorbia esula	eu	0.469	0.598	0.662	0.594	0.6	0.738	0.615	0.646	0.685	0.646
Holcus lanatus1	eu	0.512	0.673	0.698	0.576	0.542	0.633	0.434	0.521	0.619	0.722
Hypochaeris radicata2	eu	0.206	0.632	0.752	0.471	0.462	0.452	0.248	0.211	0.386	0.821
Juncus tenuis	na	0.833	0.683	0.951	0.611	0.775	0.79	0.04	0.796	0.133	0.983
Linaria vulgaris2	eu	0.949	0.903	0.952	0.924	0.805	0.951	0.946	0.962	0.952	0.773
Lythrum salicaria	eu	0.759	0.776	0.83	0.763	0.777	0.786	0.786	0.768	0.75	0.893
Medicago lupulina	eu	0.803	0.867	0.941	0.866	0.764	0.952	0.883	0.885	0.939	0.879
Melilotus albus2	eu	0.679	0.807	0.903	0.857	0.793	0.905	0.914	0.896	0.878	0.788
Phytolacca americana	na	0.575	0.476	0.923	0.368	0.855	0.655	0.286	0.514	0.548	0.951
Plantago lanceolata	eu	0.448	0.694	0.758	0.668	0.71	0.859	0.679	0.811	0.797	0.87
Plantago major	eu	0.916	0.924	0.933	0.925	0.946	0.937	0.898	0.925	0.928	0.925
Poa annua	eu	0.662	0.794	0.892	0.806	0.817	0.9	0.811	0.842	0.804	0.702
Potentilla recta	eu	0.904	0.906	0.955	0.855	0.788	0.941	0.944	0.876	0.93	0.772
Prunus serotina	na	0.697	0.702	0.981	0.656	0.969	0.592	0.422	0.956	0.984	0.991
Rhus typhina	na	0.978	0.308	0.65	0.344	0.723	0.562	0.19	0.387	0.022	1
Robinia pseudoacacia	na	0.794	0.624	0.92	0.511	0.916	0.876	0.573	0.927	0.172	0.963
Rumex acetosella	eu	0.754	0.737	0.717	0.735	0.561	0.833	0.624	0.732	0.721	0.757
Solidago canadensis2	na	0.886	0.68	0.916	0.615	0.875	0.854	0.02	0.84	0.453	0.982
Solidago gigantean	na	0.799	0.724	0.909	0.72	0.841	0.636	0.334	0.863	0.507	0.972
Sonchus oleraceus2	eu	0.717	0.72	0.843	0.703	0.756	0.718	0.692	0.778	0.706	0.771
Trifolium arvensis	eu	0.9	0.849	0.922	0.826	0.761	0.935	0.922	0.88	0.711	0.93
Trifolium dubium12	eu	0.496	0.654	0.669	0.586	0.444	0.679	0.429	0.451	0.734	0.832
Trifolium repens	eu	0.702	0.788	0.819	0.819	0.705	0.84	0.81	0.833	0.806	0.786
Verbascum thapsus	eu	0.722	0.762	0.825	0.746	0.649	0.881	0.697	0.782	0.825	0.762
Vicia sativa	eu	0.849	0.849	0.958	0.8	0.714	0.867	0.867	0.829	0.905	0.94
Acer negundo	na	0.441	0.387	0.881	0.391	0.308	0.11	0.037	0.364	0.042	0.963
Asclepsias syriaca	na	0.682	0.373	0.224	0.328	0.835	0.294	0.118	0.729	0.071	1
Aster noviibelgi	na	0.014	0.154	0.54	0.165	0.235	0.108	0.054	0.208	0.087	0.449
Bidens frondosa	na	0.926	0.698	0.951	0.656	0.884	0.667	0.467	0.889	0.674	0.926
Epilobium ciliatum	na	0.886	0.764	0.943	0.743	0.978	0.792	0.751	0.912	0.956	0.969
Helianthus tuberosus1	na	0.435	0.398	0.599	0.349	0.608	0.195	0.139	0.723	0.021	0.92
Rudbeckia laciniata	na	0.527	0.522	0.821	0.529	0.67	0.59	0.192	0.708	0.393	0.959

Table S3: Boyce index in the invaded Holarctic range for each species and strategy. Average Boyce index among 1000 replicates is provided for V_{ran} and V_{unc}.

Species	Nat	Inv	V_{all}	V_{ran}	V_{soa}	V_{unc}	V_{hie}	V_{av}	V_{avu}	V_{con}	V_{pc8}	V_{pc2}
Alliaria petiolata	eu	na	0.792	0.851	0.917	0.904	0.894	0.789	0.778	0.842	0.512	0.889
Amaranthus retroflexus	na	eu	0.921	0.786	0.955	0.865	0.986	0.933	0.597	0.981	0.927	0.896
Ambrosia artemisiifolia	na	eu	0.75	0.746	0.873	0.773	0.813	0.49	0.584	0.763	0.66	0.855
Amorpha fruticosa	na	eu	0.042	-0.234	0.065	-0.353	-0.263	-0.054	-0.636	-0.097	-0.792	-0.053
Anagallis arvensis	eu	na	0.945	0.862	0.968	0.89	0.732	0.729	0.499	0.528	0.892	0.934
Anthoxanthum odoratum	eu	na	0.861	0.831	0.922	0.804	0.841	0.878	0.784	0.812	0.883	0.909
Arabidopsis Thaliana	eu	na	0.997	0.947	0.946	0.979	0.966	0.975	0.998	0.944	0.886	0.894
Bromus sterilis	eu	na	0.116	0.392	0.8	0.452	0.607	0.271	0.674	0.64	0.506	0.572
Bromus tectorum	eu	na	0.865	0.802	0.958	0.767	0.934	0.987	0.81	0.814	0.537	0.927
Carduus nutans	eu	na	0.728	0.776	0.816	0.801	0.915	0.95	0.843	0.933	0.87	0.845
Centaurea stoebe	eu	na	0.008	0.364	0.893	0.481	0.042	0.699	0.689	0.206	0.713	-0.116
Cirsium vulgare	eu	na	0.883	0.889	0.982	0.925	0.932	0.961	0.978	0.944	0.933	0.859
Conyza canadensis	na	eu	0.936	0.827	0.879	0.851	0.963	0.874	0.716	0.938	0.847	0.981
Cytisus scoparius	eu	na	0.853	0.522	0.414	0.428	0.285	0.449	0.352	0.376	0.556	0.294
Dactylis glomerata	eu	na	0.951	0.966	0.985	0.98	0.963	0.992	0.998	0.987	0.939	0.987
Echinocystis lobata	na	eu	0.591	0.559	0.935	0.653	0.857	0.589	0.949	0.805	0.182	0.979
Erigeron annuus	na	eu	0.962	0.77	0.989	0.755	0.988	0.949	0.805	0.97	0.639	0.618
Erodium cicutarium	eu	na	0.325	0.489	0.813	0.324	0.07	0.82	0.868	0.308	0.721	0.954
Euphorbia esula	eu	na	0.847	0.814	0.862	0.832	0.829	0.901	0.975	0.846	0.798	0.838
Holcus lanatus	eu	na	0.922	0.727	0.863	0.654	0.346	0.788	0.435	0.623	0.755	0.88
Hypochaeris radicata	eu	na	0.377	0.73	0.942	0.633	0.653	0.725	0.573	0.54	0.566	0.873
Juncus tenuis	na	eu	0.529	0.381	0.806	0.402	0.776	0.259	-0.024	0.646	-0.893	0.805
Linaria vulgaris	eu	na	0.976	0.926	0.942	0.949	0.839	0.928	0.976	0.946	0.917	0.894
Lythrum salicaria	eu	na	0.641	0.658	0.641	0.61	0.608	0.007	0.814	0.307	0.3	0.888
Medicago lupulina	eu	na	0.964	0.957	0.931	0.976	0.957	0.997	0.992	0.97	0.977	0.994
Melilotus albus	eu	na	0.718	0.963	0.97	0.982	0.993	0.991	0.999	0.996	0.995	0.849
Phytolacca americana	na	eu	0.105	0.465	0.656	0.368	0.754	0.948	0.623	0.875	0.854	0.9
Plantago lanceolata	eu	na	0.042	0.823	0.813	0.849	0.901	0.984	0.971	0.961	0.937	0.981
Plantago major	eu	na	0.99	0.949	0.889	0.968	0.957	0.897	0.942	0.98	0.963	0.812
Poa annua	eu	na	0.613	0.84	0.966	0.894	0.891	0.975	0.982	0.943	0.928	0.832
Potentilla recta	eu	na	0.977	0.951	0.983	0.966	0.96	0.941	0.996	0.971	0.957	0.81
Prunus serotina	na	eu	0.98	0.494	0.966	0.429	0.857	0.553	0.44	0.592	0.952	0.855
Rhus typhina	na	eu	0.707	-0.04	0.209	0.183	0.808	-0.511	0.4	0.401	-0.676	0.916
Robinia pseudoacacia	na	eu	0.882	0.297	0.898	0.142	0.966	0.937	0.134	0.893	-0.783	0.993
Rumex acetosella	eu	na	0.624	0.597	0.597	0.719	-0.069	0.88	0.662	0.497	0.521	0.79
Solidago canadensis	na	eu	0.415	0.225	0.593	0.285	0.742	0.2	-0.843	0.532	-0.095	0.89
Solidago gigantea	na	eu	0.678	0.753	0.589	0.915	0.986	0.687	0.96	0.977	0.123	0.873
Sonchus oleraceus	eu	na	0.875	0.878	0.931	0.884	0.969	0.959	0.965	0.971	0.62	0.783
Trifolium arvensis	eu	na	0.817	0.875	0.891	0.871	0.85	0.922	0.942	0.635	0.98	0.963
Trifolium dubium	eu	na	0.143	0.651	0.901	0.673	0.601	0.703	0.749	0.8	0.764	0.882
Trifolium repens	eu	na	0.804	0.855	0.943	0.936	0.83	0.957	0.97	0.979	0.97	0.868
Verbascum thapsus	eu	na	0.879	0.924	0.975	0.97	0.99	0.966	0.989	0.991	0.971	0.897
Vicia sativa	eu	na	0.911	0.891	0.889	0.861	0.607	0.936	0.994	0.969	0.895	0.885
Acer negundo	na	eu	-0.388	-0.196	0.937	-0.018	-0.717	-0.812	-0.468	-0.149	-0.873	0.92
Asclepsias syriaca	na	eu	0.516	0.296	0.48	0.25	0.548	-0.045	0.29	0.37	-0.584	0.84
Aster noviibelgi	na	eu	-0.744	0.242	-0.122	0.264	0.67	0.187	-0.15	0.815	-0.372	0.638
Bidens frondosa	na	eu	0.954	0.567	0.749	0.576	0.963	0.587	0.515	0.859	0.382	0.933
Epilobium ciliatum	na	eu	0.404	0.249	0.802	0.23	0.887	0.55	0.474	0.486	0.347	0.977
Helianthus tuberosus	na	eu	0.911	0.253	0.367	0.342	0.934	-0.55	0.173	0.981	-0.887	0.937
Rudbeckia laciniata	na	eu	0.786	0.533	0.861	0.73	0.818	0.844	0.789	0.777	-0.033	0.836

Table S4: Sensitivity in the invaded Australian range for each species and strategy. Average Sensitivity among 1000 replicates is provided for V_{ran} and V_{unc}.

Species	Nat	V_{all}	V_{ran}	V_{soa}	V_{unc}	V_{hie}	V_{av}	V_{avu}	V_{con}	V_{pc8}	V_{pc2}
Amaranthus retroflexus	na	0.038	0.36	0.283	0.432	0.472	0.66	0.509	0.415	0.396	0.962
Ambrosia artemisiifolia	na	0.887	0.703	0.962	0.827	0.811	0.925	0.849	0.868	0.226	0.925
Anagallis arvensis	eu	1	0.983	1	0.988	1	0.768	0.951	0.996	1	0.996
Anthoxanthum odoratum	eu	0.193	0.793	0.924	0.861	0.959	0.912	0.918	0.971	0.959	0.702
Arabidopsis Thaliana	eu	0.667	0.901	1	0.955	1	1	1	1	1	1
Bromus sterilis	eu	0.829	0.893	0.943	0.871	0.943	0.914	0.857	0.886	0.943	0.857
Bromus tectorum	eu	0.964	0.961	0.929	0.964	0.964	0.964	0.964	0.964	0.964	1
Carduus nutans	eu	0.983	0.986	0.983	0.991	1	0.966	1	0.983	1	0.983
Cirsium vulgare	eu	0.29	0.648	0.753	0.658	0.637	0.876	0.736	0.729	0.729	0.525
Conyza canadensis	na	1	0.979	0.941	0.974	1	0.912	0.961	0.98	0.961	0.99
Cytisus scoparius	eu	0.948	0.95	0.99	0.943	0.938	0.969	0.918	1	0.979	0.959
Dactylis glomerata	eu	0.984	0.983	0.984	0.986	0.973	0.968	0.995	0.995	0.978	0.941
Erodium cicutarium	eu	0.637	0.66	0.847	0.715	0.738	0.732	0.818	0.769	0.826	0.981
Holcus lanatus	eu	0.839	0.936	0.963	0.895	0.935	0.912	0.834	0.935	0.945	0.88
Hypochaeris radicata	eu	0.572	0.833	0.9	0.709	0.714	0.912	0.646	0.714	0.888	0.844
Juncus tenuis	na	0.84	0.879	0.98	0.906	1	0.88	0.82	0.96	0.08	0.94
Linaria vulgaris	eu	0	0.599	0.571	0.695	0.714	0.857	0.786	0.643	0.5	0
Lythrum salicaria	eu	0.919	0.891	0.965	0.875	0.931	0.988	0.879	0.936	0.942	0.85
Medicago lupulina	eu	0.754	0.868	0.905	0.855	0.77	0.929	0.802	0.897	0.937	0.873
Melilotus albus	eu	0.476	0.663	0.634	0.752	0.585	0.963	0.927	0.78	0.829	0.268
Phytolacca americana	na	1	0.991	1	0.991	1	1	1	1	1	1
Plantago lanceolata	eu	0.788	0.741	0.9	0.806	0.83	0.934	0.88	0.927	0.915	0.857
Plantago major	eu	0.892	0.932	0.954	0.944	0.923	0.915	0.908	0.931	0.938	0.838
Poa annua	eu	0.852	0.895	0.916	0.895	0.905	0.935	0.932	0.905	0.897	0.437
Potentilla recta	eu	0.075	0.725	0.9	0.87	1	1	1	1	1	0.925
Robinia pseudoacacia	na	0.789	0.589	0.8	0.563	0.822	0.633	0.811	0.833	0.389	0.933
Rumex acetosella	eu	0.713	0.9	0.92	0.858	0.816	0.77	0.828	0.897	0.9	0.686
Solidago canadensis	na	0.154	0.143	0.231	0.235	0.096	0.673	0.269	0.269	0	0.5
Sonchus oleraceus	eu	0.533	0.619	0.83	0.581	0.632	0.636	0.525	0.68	0.728	0.475
Trifolium arvensis	eu	0.762	0.798	0.802	0.77	0.741	0.935	0.75	0.873	0.793	0.716
Trifolium dubium	eu	0.777	0.827	0.964	0.824	0.853	0.79	0.853	0.884	0.768	0.821
Trifolium repens	eu	0.845	0.913	0.942	0.876	0.801	0.907	0.872	0.929	0.863	0.819
Verbascum thapsus	eu	0.826	0.909	0.959	0.883	0.835	0.959	0.851	0.868	0.95	0.554
Vicia sativa	eu	0.909	0.893	0.984	0.863	0.807	0.945	0.874	0.882	0.941	0.945
Acer negundo	na	0.952	0.885	0.952	0.85	0.952	0.762	0.429	1	0.143	1
Aster noviibelgi	na	0	0.117	0.85	0.125	0.2	0.1	0.1	0.15	0	0.15
Epilobium ciliatum	na	0.945	0.769	0.945	0.688	0.982	0.364	0.182	0.527	0.582	0.891
Helianthus tuberosus	na	0.636	0.262	0.182	0.216	0.364	0.455	0.364	0.091	0	0.818

Table S5: Boyce index in the invaded Australia range for each species and strategy. Average Boyce index among 1000 replicates is provided for V_{ran} and V_{unc}.

Species	**Nat**	**V_{all}**	**V_{ran}**	**V_{soa}**	**V_{unc}**	**V_{hie}**	**V_{av}**	**V_{avu}**	**V_{con}**	**V_{pc8}**	**V_{pc2}**
Amaranthus retroflexus	na	-0.585	0.394	-0.154	0.551	0.766	0.802	0.541	0.643	0.513	0.659
Ambrosia artemisiifolia	na	0.669	0.774	0.92	0.814	0.727	0.849	0.887	0.852	0.056	0.919
Anagallis arvensis	eu	0.907	0.736	0.807	0.838	0.911	0.913	0.918	0.902	0.623	0.526
Anthoxanthum odoratum	eu	0.823	0.914	0.972	0.91	0.926	0.92	0.873	0.925	0.968	0.949
Arabidopsis thaliana	eu	0.324	0.676	0.755	0.728	0.627	0.694	0.722	0.599	0.874	0.771
Bromus sterilis	eu	0.656	0.607	0.643	0.622	0.491	0.664	0.626	0.584	0.622	0.608
Bromus tectorum	eu	0.546	0.655	0.851	0.609	0.665	0.586	0.513	0.669	0.636	0.926
Carduus nutans	eu	0.961	0.819	0.869	0.878	0.917	0.666	0.809	0.821	0.841	0.907
Cirsium vulgare	eu	0.829	0.908	0.953	0.894	0.943	0.938	0.979	0.978	0.991	0.921
Conyza canadensis	na	0.831	0.392	0.173	0.387	0.309	-0.343	0.617	0.613	0.139	0.114
Cytisus scoparius	eu	0.795	0.842	0.565	0.871	0.865	0.918	0.892	0.914	0.879	0.9
Dactylis glomerata	eu	0.912	0.853	0.896	0.82	0.794	0.619	0.819	0.843	0.818	0.976
Erodium cicutarium	eu	0.99	0.915	0.958	0.897	0.925	0.891	0.963	0.898	0.978	0.906
Holcus lanatus	eu	0.648	0.702	0.901	0.539	0.466	0.894	0.813	0.661	0.768	0.954
Hypochaeris radicata	eu	0.887	0.814	0.867	0.662	0.358	0.932	0.916	0.703	0.716	0.932
Juncus tenuis	na	0.838	0.855	0.894	0.847	0.776	0.843	0.964	0.794	-0.638	0.545
Linaria vulgaris	eu	0.644	0.5	0.554	0.575	0.312	0.665	0.748	0.454	0.396	-0.378
Lythrum salicaria	eu	0.872	0.844	0.868	0.839	0.642	0.901	0.962	0.576	0.952	0.897
Medicago lupulina	eu	0.906	0.887	0.884	0.85	0.698	0.767	0.907	0.852	0.913	0.958
Melilotus albus	eu	0.861	0.863	0.948	0.892	0.927	0.704	0.863	0.949	0.901	0.671
Phytolacca americana	na	0.857	0.825	0.847	0.801	0.746	0.636	0.839	0.651	0.935	0.812
Plantago lanceolata	eu	0.97	0.868	0.891	0.841	0.827	0.699	0.828	0.969	0.951	0.948
Plantago major	eu	0.856	0.834	0.92	0.882	0.951	-0.022	0.921	0.967	0.885	0.892
Poa annua	eu	0.98	0.913	0.877	0.919	0.929	0.76	0.892	0.964	0.961	0.918
Potentilla recta	eu	0.969	0.803	0.854	0.772	0.856	0.695	0.667	0.75	0.819	0.946
Robinia pseudoacacia	na	0.952	0.798	0.949	0.736	0.923	0.986	0.852	0.969	0.767	0.964
Rumex acetosella	eu	0.937	0.933	0.954	0.925	0.97	0.977	0.977	0.948	0.929	0.928
Solidago canadensis	na	-0.341	-0.108	0.156	0.129	-0.048	0.328	0.548	0.604	0.015	0.868
Sonchus oleraceus	eu	0.937	0.906	0.876	0.951	0.914	0.902	0.962	0.939	0.987	0.941
Trifolium arvensis	eu	0.931	0.922	0.917	0.918	0.96	0.92	0.967	0.969	0.987	0.956
Trifolium dubium	eu	0.984	0.934	0.97	0.944	0.949	0.969	0.941	0.986	0.97	0.939
Trifolium repens	eu	0.957	0.909	0.952	0.887	0.839	0.806	0.931	0.929	0.865	0.934
Verbascum thapsus	eu	0.863	0.857	0.888	0.901	0.856	0.807	0.928	0.863	0.866	0.972
Vicia sativa	eu	0.902	0.925	0.988	0.911	0.758	0.952	0.863	0.956	0.899	0.978
Acer negundo	na	0.815	0.788	0.832	0.823	0.878	0.833	0.752	0.869	0.019	0.635
Aster noviibelgi	na	-0.357	-0.218	0.214	-0.351	-0.246	-0.571	-0.529	-0.192	-0.379	0.115
Epilobium ciliatum	na	0.924	0.776	0.907	0.661	0.844	0.149	0.175	0.918	0.579	0.466
Helianthus tuberosus	na	0.461	-0.031	0.022	-0.181	-0.188	0.049	-0.214	-0.29	-0.636	0.608

Table S6: Coefficients of Generalized linear models built on traits values to explain transferability in the invaded range for non-shifting species.

Invaded range: Holarctic *Strategy : Vsoa* *Evaluation : Se*

	Estimate	Std. Error	t value	Pr(>|t|)	.
(Intercept)	6.23E-01	5.48E-02	11.361	3.05E-14	***
Native niche breadth	1.17E-04	2.70E-05	4.34E+00	9.12E-05	***

Invaded range: Holarctic *Strategy : Vsoa* *Evaluation : B*

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.69687	0.05145	13.545	2.00E-16	***
Selfing:yes	0.19863	0.06376	3.115	0.00335	**

Invaded range: Holarctic *Strategy : Vpc2* *Evaluation : Se*

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	6.98E-01	5.38E-02	12.981	9.78E-16	***
Sp.Origin:NA	8.13E-02	3.45E-02	2.358	0.0235	*
Native niche breadth	6.89E-05	2.37E-05	2.912	0.00591	**
Height	5.80E-05	3.04E-05	1.905	0.06411	.

Invaded range: Holarctic *Strategy : Vpc2* *Evaluation : B*

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	8.27E-01	3.37E-02	24.522	2.00E-16	***
Native niche breadth	4.36E-05	1.60E-05	2.721	0.00959	**
Ploidy: >2	-5.92E-02	2.18E-02	-2.712	0.00982	**

Invaded range: Australia *Strategy : Vsoa* *Evaluation : Se*

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	7.52E-01	1.06E-01	7.126	9.39E-08	***
Native niche breadth	8.79E-05	4.67E-05	1.885	0.0699	.
Sp.Origin:NA	-1.21E-01	6.53E-02	-1.848	0.0752	.

Invaded range: Australia *Strategy : Vsoa* *Evaluation : Se*

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.86714	0.05663	15.311	1.98E-15	***
Sp.Origin:NA	-0.30674	0.09972	-3.076	0.00454	**

Invaded range: Australia *Strategy : Vpc2* *Evaluation : Se*

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.91914	0.04808	19.12	<2e-16	***
Repro.veg	-0.12267	0.06492	-1.89	0.0688	.

Invaded range: Australia *Strategy : Vpc2* *Evaluation : B*

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	0.89286	0.04154	21.495	2.00E-16	***
Sp.Origin:NA	-0.30916	0.07314	-4.227	0.000215	***

CHAPTER 1.4

Native climate niche predicts the introduction success of alien mammals

Olivier Broennimann[1], Blaise Petitpierre[1], Sven Bacher[2], Antoine Guisan[1]

[1] Department of Ecology & Evolution, University of Lausanne, 1015 Lausanne, Switzerland

[2] University of Fribourg, Department of Biology, Unit of Ecology & Evolution, Chemin du Musée 10, 1700 Fribourg, Switzerland

This paper is in preparation

My contribution to the project: I participated to the data collection, developed and ran the niche modeling analyses and co-wrote the initial draft of the paper with OB. Both fist authors contributed equally.

ABSTRACT

Understanding the factors behind the success of introduction of invasive species is crucial for a better anticipation. While information about successful invasive are abundantly reported, failure of introduction are scarcely reported but crucial to untangle the numerous possible factors making a naturalized species successful. Among numerous possible factors, climate matching between native niches and sites of introductions can play a significant role. We tested native climate niche matching for 168 mammal species for which introduction status was known at worldwide scale. We show that successful introduced mammals tend to be introduced in similar climate as their native distribution. Combining niche matching to other traits affecting invasion success should help to understand the first step allowing species to become invasive.

INTRODUCTION

The establishment and spread of introduced species is now recognized as a major risk worldwide. This risk is likely to further increase with global change as a result of greater volumes of transport and trade that ultimately will translate into increasing rates at which novel species are introduced to new locations (e.g. Mack *et al.*, 2000). It is thus important to anticipate risk of invasion and design appropriate conservation measures to reduce the amount of introduction events (Simberloff, 2009).

Progress has been made in recent years to identify the factors responsible for the success of establishment of species (e.g. Blackburn and Duncan, 2001, Kolar and Lodge, 2001, Kolar and Lodge, 2002, Duncan *et al.*, 2003; Hayes and Barry, 2008 for a review), but it was shown that failures and success highly depend on the idiosyncrasies of the release event and the characteristics of the recipient community (Blackburn and Duncan, 2001, Duncan *et al.*, 2003). Moreover, this type of analysis is difficult because i) the set of species analyzed is never random because some species were introduced intentionally by humans, and ii) introductions happen disproportionately in some areas (e.g. in areas with high human activities or close to major transportation hubs). As a result, individual introductions cannot be considered as statistically independent (Blackburn and Duncan, 2001, Simons, 2003, Sol *et al.*, 2008). The appropriate way to test for factors responsible for establishment success is thus to use statistical approaches that contrast failures and success of introduction while simultaneously controlling for potentially confounding factors. Without proper testing, confounding factors such as taxonomy or biogeographical provenance of species can either obscure the relationship between the success of establishment and the variable of interest (Blackburn and Duncan, 2001).

Because introduction success largely depends on the taxonomic group studied (e.g. Hayes and Barry, 2008, Allen *et al.*, 2013), here we concentrate on invasive mammals. We chose mammals as a study group because it contributes a large amount of invasive species in the world (Myers, 1986, Long, 2003) and, because of their proximity to human,

the failures and successes of their introduction to new areas are well documented (e.g. Myers, 1986, Long, 2003). Among the studies contrasting the success and failures of introductions for mammals (i.e. we do not mention studies investigating only situations of introduction success), the factors best explaining the introduction success of mammals vary strongly between areas: at worldwide scale they include size of the brain and introductions effort (Sol *et al.*, 2008), in Florida only the number of already introduced invasive species was significant (Allen *et al.*, 2013), in Australia only climatic suitability and introduction effort were retained by models (Forsyth *et al.*, 2004) and in New Zealand it was the introduction effort (Forsyth and Duncan, 2001). To our knowledge, no study investigates the importance of climate suitability at the worldwide scale.

Ecological niche models (ENM) based on the climatic niche are now routinely used to predict areas at risk of invasion by exotic species (e.g. Pheloung *et al.*, 1999, Thuiller *et al.*, 2005, Jimenez-Valverde *et al.*, 2011, Vaclavik and Meentemeyer, 2012 but see Guisan *et al.*, 2013). The practice relies on the assumption that species occupy the same macroclimatic niche in their native and invasive range, i.e. that the macroclimatic niche is conserved in space and time. Models can be calibrated on the native range and then projected to the invaded range. Petitpierre et al. (2012) showed that this assumption is valid for the majority of Holarctic plant invaders in analog climates. However, some invasive species occupy climatic conditions in their invaded range that differ from their native niche (e.g. Broennimann *et al.*, 2007, Fitzpatrick *et al.*, 2007, Rödder and Lötters, 2009, Medley, 2010, Gallagher *et al.*, 2010, Kolbe *et al.*, 2012, Petersen, 2013). Reasons for such climatic niche shifts vary idiosyncratically between species but include dispersal limitation in the native range (Guisan *et al.*, 2014), competition and biotic interactions release (Wolfe, 2002, Prenter *et al.*, 2004,) and evolutionary processes such as reallocation of defenses towards growth (Blossey and Notzold, 1995, Maron *et al.*, 2004). For these species the climate matching approach used in ENM projections cannot be used to assess the full extent of invasions. However, the question remains whether ENM projections are useful to identify areas at risk of introduction and establishment of newly or not-yet-introduced invasive species, i.e. before ecological and evolutionary processes would alter the niche of the species. Some case studies showed that even niche shifting species had their historical introduction in similar climatic conditions as the environment (e.g. spotted knapweed, Broennimann *et al.*, 2007, Broennimann *et al.*, 2013) but to our knowledge, this has not been formally tested with a large set of species.

Here we test whether the success and failures of worldwide introductions of mammals can be explained by the climatic conditions at the site of introduction as predicted by models of the native niche. For this test we use both ENM and ordination techniques. We predict that the more the site of introduction is located at the center of the native niche, the more the introduction will be successful. We further discuss how the results integrate in the general discussion of niche conservatism during biological invasion and contrast the perspective on the predictability of introductions vs. predicting the full invaded range.

METHODS

General framework of analyses

Two approaches were followed to measure de degree to which introductions occurred in macroclimatic conditions corresponding to the core of the native niche (Fig. 1):

ENM approach: species distribution models for each species were calibrated in their native range and projected to the rest of the world, including areas of introductions. We retrieved the predicted habitat suitability (or the maximum suitability for polygons and lines) at the introduction sites (HS_{intro}) and measured its quantile within the distribution of values of habitat suitability of pixels occupied by the species in the native range (HS_{NATIVE}). We refer to this quantile the innerness of the introduction to the native niche as measured by ENM suitability as $INNER_{HS}$ in the rest of the manuscript.

Ordination approach: we calibrated a principal component analysis maximizing the macroclimatic variation over all the pixels of the study area and calculated the density of native occurrences within the environmental space depicted by the two first axes (i.e. PCA-env approach in Broennimann *et al.*, 2012). We retrieved the densities of occurrence (or the maximum density for polygons and lines) corresponding to the conditions at the site of introduction ($dens_{intro}$) and measured its quantile within the distribution of density values in the native range ($dens_{NATIVE}$). We refer to this quantile the innerness of the introduction to the native niche as measured by species density in the PCA space as $INNER_{dens}$ in the rest of the manuscript.

Introductions were considered to take place inside of the native niche when as $INNER_{HS}$ or $INNER_{dens}$ were superior to 0.05. Although arbitrary, this threshold corresponds to the 95^{th} percentile of the distribution, a typical threshold used in statistical tests for significance. Moreover, it typically corresponds to the ratio of presences correctly predicted by ENM (i.e. sensitivity) when the latter is binarized using a threshold optimizing for both presences and absences correctly predicted. We then modeled the statistical relationship between the success of introductions and INNER using binomial mixed effect models with phylogenetic relatedness as a random factor.

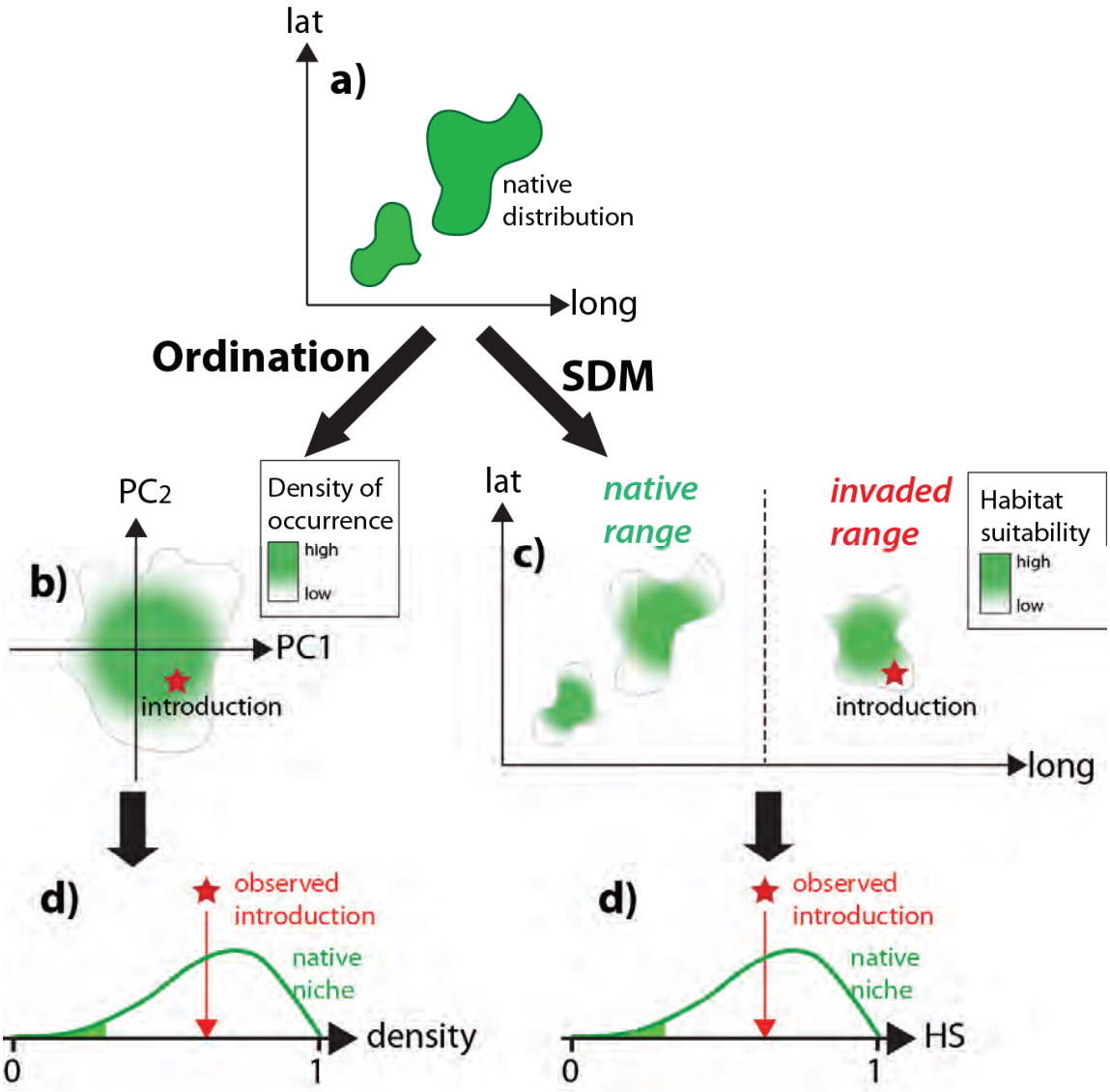


Figure 1: Schematic representation of the two types of analyses of belonging of introduction to the native niche. a) native distribution of the species. b) density of occurrence of the species along the axes of a PCA (see Broennimann et al. 2012). c) SDMs' predictions in the geographic space. d) introduction position relatively to the native niche: suitability/density at the introduction site is compared with the distribution of suitability/density covered by the native species distribution.

Introductions

The introductions referenced in John Long's book (Long, 2003) on introduced mammals of the world were screened and georeferenced. Depending on the precision of the description of introductions, points, polygons or lines were digitized in ArcMap. Polygons with an area larger than 10'000 pixels were not included in the analysis. We recorded the success of introduction (i.e. extinct, declining, established, expanding), the type of introduction (i.e. intentional or unintentional), as well as the type of release (i.e. deliberate or accidental). Note that reintroductions where not used for the analysis.

Native distributions

The native distributions of each mammal's species for which introduction data were available were extracted from the 2008 IUCN Red List (IUCN 2008), which is the world's most comprehensive inventory of the global conservation status for mammal species. Shapefiles were rasterized at the same resolution as climate data (i.e. 10 Arc minutes ~ 16 km.).

Climate data

Seven climate variables from the Worldclim dataset (Table 1, Hijmans *et al.*, 2005) and one aridity variable (available on on the CGIAR-CSI website http://www.cgiar-csi.orgm, Zomer *et al.*, 2008) were used to assess species native bioclimatic niche. These variables were chosen to limit the colinearity and to be ecologically relevant for the survival of mammal's species. Resolution of the environmental grids was 10 Arc minutes (about 16 km).

Table 1: Variables used for species distribution models

Abbreviation	**Description**
ai	Aridity index
td	Mean Diurnal Range (Mean of monthly (max temp - min temp))
ts	Temperature seasonality
twetq	Mean temperature of the warmest quarter
tcoldq	Mean temperature of the coldest quarter
pwetq	Precipitation of the warmest quarter
pdryq	Precipitation of the driest quarter
pwarmq	Precipitation of the warmest quarter

Study area

The choice of the study areas (also called background areas) used to calibrate ENM may have severe effects on spatial predictions (Barve *et al.*, 2011). It is recommended to delimit the study areas in regards to species historical biogeography (Barve *et al.*, 2011). In this study, three approaches to delimit study areas were used, from coarser to finer scale : i) all pixels worldwide at 16km resolution, ii) all pixels of the zoogeographical region(s) where the species occurs in its native distribution and iii) only the landuse categorie(s) within the zoogeographical region(s) where species occurs. The worldwide approach is too coarse, affected by biases due to species dispersal limitation but allows projecting SDM on the whole world as no place is located in non-analog climates (see below). We used the definition and cartography from Cox (2001) to delimit species zoogeographic regions. These latter areas are characterized by containing endemic families, subfamilies and tribes, with very high generic and species endemism. They include the main biogeographic barriers and can be considered as a coarse but consistent biogeographic extents), at least for Mammals with a widespread distribution (Cox, 2001).

As all types of climates may not be represented within each region, non-analog climates may appear when projecting ENM on the rest of the world. Eventually, including a finer dimension such as landuse can be particularly relevant for mammal species breading on particular type of landuse categories. However, the amount of non-analog climates increases when ENM are projected to the rest of the world increases with refining the study area.

ENM

Native species occurrences were related to environmental predictors using three different modeling techniques averaged in an ENSEMBLE forecasting approach (Araujo and New, 2007). Generalized linear models with polynomial quadratic coefficients preceded by a stepwise selection based on the Bayesian Information Criteria (McCullagh and Nelder, 1989), Boosted regression tree (with 5000 trees and 7 degrees of interaction depth, Friedman, 2001) and a MAXENT, a technique based on maximum entropy (with default parameters, Phillips *et al.*, 2006). The BIOMOD package (Thuiller et al., 2009) was used to calibrate GLM and GBM whereas MAXENT was done with dismo package (Hijmans, 2010) on the R software (R, 2005). Occurences were combined with 10’000 pseudoabsences randomly sampled throughout the study area and weighted such as both occurrences and pseudoabsences shared the same prevalence. The whole dataset was split into 70 % for calibration while the remaining 30 % were kept for ENM evaluation on pseudo-independent data. Variables contribution was assessed through randomization (see the documentation of BIOMOD). Suitability and variables contribution were averaged between each technique. The resulting averaged suitabilities were evaluated with the Area Under the Curve of a Receiver Operating Characteristics (AUC, Zweig and Campbell, 1993) and True Skill Statistics (TSS) known to be more robust to differences among species prevalence and variation of the study area (Allouche *et al.*, 2006). For each species, ENM were then projected such as we obtained worldwide potential distribution, including all introductions. When introductions sites were covered by several pixels, we attributed the maximum suitability index to the introduction. We excluded pixels were the climate that was non-analog to the species native range for two reasons. The interpretation of introductions occurring in non-analog climate remains highly speculative because climates are considered as “non-accessible” in the native range, and could be due either to preadaptation or to niche shift. Second, ENM predictions towards non-analog climates have been shown to be highly hazardous and difficult to interpret (Fitzpatrick and Hargrove, 2009). Non analog climates were considered as soon as one of the predictors outreached the range cover in the native study area. Thus, ENM predictions were attributed to every introductions site.

PCA-env

For each species, PCA were calibrated along the environmental predictors covered by the whole native study area. It has been shown to be an appropriate approach to estimate niche overlap in Holarctic environment (Broennimann *et al.*, 2012). Environmental space depicted by the two first axes of the PCA was then gridded (resolution = 100 by 100) and species occurrences were then converted into species smoothed densities using a simple kernel function. The procedure was done using the same code as Broennimann *et al.*, 2012 in R. For each species, introductions were projected to this environmental space and compared the density of the native niche in similar conditions.

Mixed effect models

To analyze the relationship between introduction success and the niche innerness, we fitted generalized linear mixed-effect models using the MCMCglmm library (Hadfield, 2010) in R. We used a sample of the phylogenetic tree of mammals (Bininda-Emonds *et al.*, 2007) including 95 of our initial species to account for phylogenetic relatedness between species in the model. Both $INNER_{dens}$ and $INNER_{dens}$ using the three calibration areas were tested as an explanatory variable. For clarity reason, we only present results with niche calibrated at the scale of zoogeographical regions.

RESULTS

ENM Evaluation

On the average, ENMs had excellent AUC and TSS scores in their native range with the three calibration approaches (Fig. 2). AUC values varied between 0.810 and 1 with a median of 0.990 whereas TSS varied between 0.478 and 1 with an average of 0.922. Evaluation indices slightly decreased with refined calibration extents (Fig. 2).

The median inertia explained by the first and second axes of the PCA was 51% and 26% repectively varying between 38% and 67% for axis 1, 20% to 37% for axis 2.

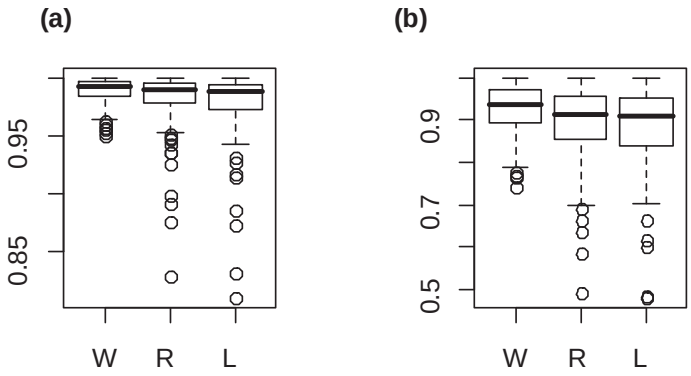


Figure 2: Evaluations of the SDM with AUC (a) and TSS (b) when calibrated at the worldwide scale (W), zoogeographical regions scale (R) and landuse scale (L).

Success of introductions

We retained 701 introductions for 168 species (Fig. 3). The median size of the native area was 3792 pixels varying between 27 and 298642 (Fig. S1). The median number of introduction per species was 2 and varied between 1 and 23 (Fig. S2). We inventoried 69 failed introductions (64 extinctions and 5 declining populations), 550 successful introductions (197 established and 353 expanding populations) and 82 for which status was impossible to establish (59 eradicated and 23 for which no information was available). There was a slight but significant positive correlation effect of the size of the native area and the number of introduction, even when outliers (defined as the 2.5% at each of the two extremities of the area distribution) were removed (Fig. S3). There was no significant differences in native area size between failed and successful introductions (t-test P = 0.466). Six introductions were removed when filtering model calibration by zoogeographical regions and by realms because they fell in non-analog climates.

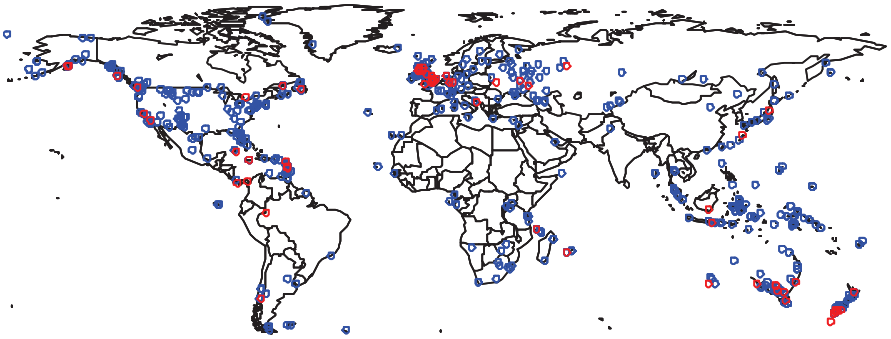


Figure 3: Distribution of the successful (blue) and failed (red) introduction over the world.

Number of introductions belonging to the native niche

There were significant effects of calibration extent and reported introduction success on the predictability of introductions (Table 3, Fig. 3). 42 % to 51% of successful introductions belonged to the native niche as depicted by ENM (Table 3), contrasting with only 26 % to 35 % for the failed introductions. The PCA analysis supported the same trend, without any effect of the calibration extent. 64% to 66% of the successful introductions belonged to the native niche contrasting with 41% to 47% of the failed introductions.

Table 2: proportion of introductions belonging to the native niche (i.e. > 5th quantile of HS_{native}) following SDM/PCA calibration (worldwide, zoogeographical regions or landuse) and reported success of introduction (failed, successful or no information)

	World (SDM/PCA)	Z. regions (SDM/PCA)	Landuse (SDM/PCA)
Failed	32% / 41%	35% / 46%	32% / 48%
Successful	42% / 64%	51%/ 66%	50% / 66%
NA	26% / 41%	32% / 47%	35% / 48%

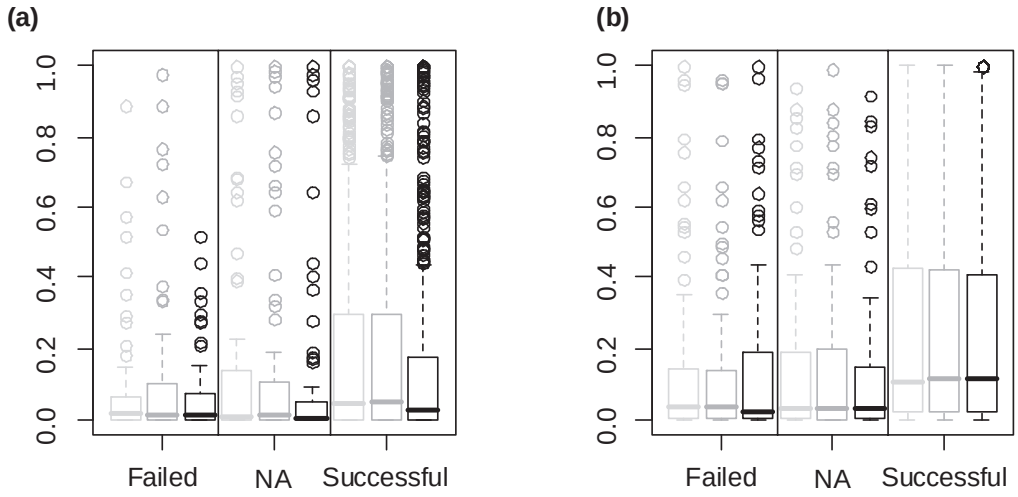


Figure 4: Boxplot of the HS_{intro} and dens $_{intro}$ relatively to the native distribution, in regard to introduction success (Failed, Succesful or NA) and calibration extent (world in grey, realms in darkgrey and landuse in black.

Species level

Between 66 and 94 species (39% and 56% respectively) had more than half of their introductions belonging to their native niche depending on the technique (SDM or PCA) and the extent used to assess the niche (Table 4). Native range size, regions and classes of landuse covered in the native range did not show any significant effect on the proportion of introduction belonging to the native niche (ANOVA test for native range size and Chi-square test for classes of landuse and regions).

Table 3: Number of species with respectively 100%, between 51% and 99%, Between 1% and 50% and none of their introduction predicted by SDM/PCA following different calibration approaches (zoogeographical regions or landuse).

	World (SDM/PCA)	**Z. Regions (SDM/PCA)**	**Landuse (SDM/PCA)**
100%	51 / 62	50 / 64	50 / 64
between 51% and 99%	15 / 30	26 / 30	26 / 60
Between 1% and 50%	46 / 34	31 / 34	31 / 34
0%	56 / 38	59 / 38	59 / 38

Mixed effect models

The mixed effect models predicting the success of introduction as response to native niche innerness show a significant of native climatic niche matching (Fig. 5; Table 5 and 6). The response was fitted as a binomial response to the square root of niche innerness as a fixed factor, while correcting for the phylogenetic relatedness between random factors.

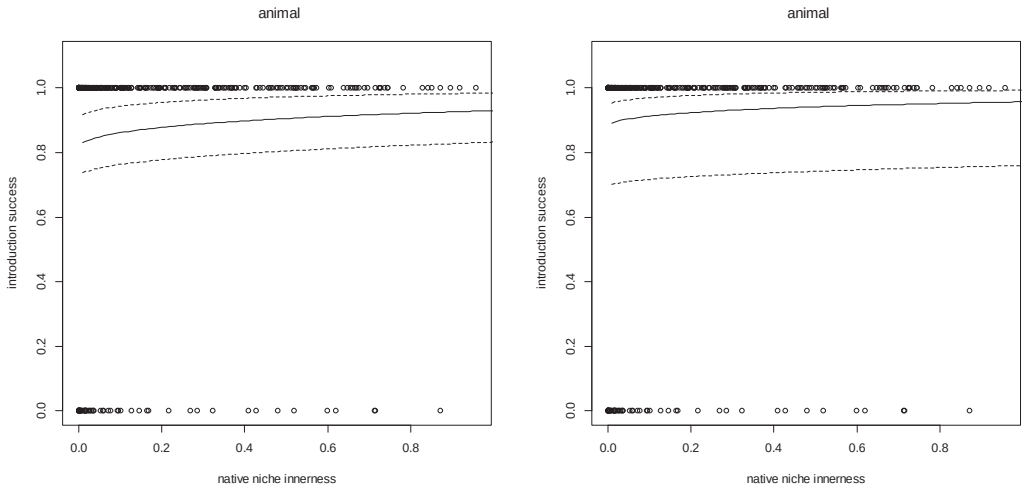


Figure 5: relationship between introduction success and native niche innerness measure with ordinations (a) and SDMs (b) calibrated on zoogeographical regions

Table 4: summary of the mixed model for $INNER_{dens}$ (calibration on zoogeographical regions)

type	factor	post.mean	Lower 95% CI	Upper 95% CI	Effective sample	p-value (MCMC)
fixed	(intersept) sqrt(dens)	1.4769 1.1099	0.9668 0.6363	2.1955 1.9487	4.986 3.114	<0.001 <0.001
random	Phylogeny Unit	0.1829 0.00753	0.0002092 0.000311	0.3672 0.02467	9.7661 4.68	

Table 5: summary of the mixed model for $INNER_{HS}$ (calibration on zoogeographical regions)

type	factor	post.mean	Lower 95% CI	Upper 95% CI	Effective sample	p-value (MCMC)
fixed	(intersept) sqrt(dens)	1.9985 1.1081	0.8264 0.3262	2.7923 2.1526	25.49 18.01	<0.001 0.006
random	Phylogeny Unit	2.368 0.08491	0.2546 0.001403	4.46 0.3443	18.46 17.72	

DISCUSSION

For the first time, climate matching was tested not only on introduction success but also on introduction failures of introduced mammals at worldwide scale. This study reveals that failed mammals introductions correspond on average to sites with conditions further away from the optimal native niche (i.e. having lower average native niche innerness) than successful introductions. However, only 50 to 60 % of the successful introductions belonged to their native niche, suggesting that climate matching only partially explains introductions success in mammals. We discuss the ability and the limit of species distribution predictions based on the bioclimatic niche in the light to other factors that could explain invasion success.

Niche conservatism in biological invasions: from predicting successful introductions to full invaded ranges

Niche conservatism is a debated topic for mammals (Dormann *et al.*, 2010, Cooper *et al.*, 2010). One can reasonably hypothesize that the decoupling between the organism's physiology and its bioclimatic environments due to endothermy also decouples the realized niche from the fundamental niche of mammals. In other words, mammals' distributions may reflect competitions more than their tolerance to the environment. Thus predictions of their potential distributions through the use of bioclimatic models could be biased (Dormann *et al.*, 2010). In this regard, biological invasions can be seen as worldwide and unplanned experimental opportunity to test niche conservatism and the ability of species to adapt to novel conditions (Sax *et al.*, 2007). Our results support that realized niche conservatism is a partial but important component of successful mammals' introductions. They confirm a previous study focused on 40 introduced mammals in Australia, which show a significant contribution of climate matching (combined with introduction efforts, ability of migration and the size of the range oversea, Forthys *et al*. 2004). These results were however not supported by a study investigating introduction success for 25 mammal species in Florida, where only the number of other non-native species already naturalized in the introduction area had a significant explanation (Allen *et al.* 2013). In such subtropical conditions, the role of climate in driving species distribution may be less important than under more stressful conditions (Normand *et al.*, 2009, Araújo *et al.*, 2013).

To our knowledge, reptiles and amphibians are the only taxonomical groups for which a worldwide test of whether it matters for a species to be introduced in climatic conditions defined by its native niche as been performed (Bomford *et al.*, 2009) using a simplistic envelope technique (Climate software; Pheloung, 1996). For mammals, such an investigation still remains to be done. The only worldwide study available on mammals (Sol *et al*, 2008) showed no biome conservatism, which was included as a proxy for native climate matching. We feel however that this binary variable is not specific and discriminating enough to properly investigate this hypothesis. For instance, it was shown

that biomes where South African species occur in their native range are not reliable descriptors the biomes that they can invade elsewhere because it depends strongly on how biomes are defined in different parts of the world (Thuiller *et al.*, 2005).

Our results show that although significant, native climatic niche is only a partial component for species introduction success. It could be associated – e.g. in invasive species risk assessments - with other environmental information (e.g. microhabitat data on landuse or topography) or species traits. Coupling species traits to ENMs allows a wider understanding of mechanisms behind the success of invasive species, with potential to improve climate matching analyses and spatial predictions (Moles *et al.*, 2008, Hulme and Barrett, 2013). Until now, among the numerous traits tested, only brain size relative to body mass and introduction effort were found as the only significant factors explaining invasion success at the worldwide scale (Sol *et al.*, 2008a). Brain size is intimately linked to behavioral plasticity and the ability of learning, two crucial skills for a successful adaptation to novel environments (Sutter and Kawecki, 2009, Chapple *et al.*, 2012). Combining these latter traits to the native climatic niche could produce more powerful predictions of invasions success. Such predictions are particularly desirable as prevention, done for example through pre-border risk assessment (Koop *et al.*, 2012), is much more efficient than eradication (Leung *et al.*, 2002).

Introduction is only the first stage of the naturalization and then possibly the invasion (Richardson *et al.*, 2000). This study supports that climate is an important driver of species invasion (Pyšek *et al.*, 2010). It is then relevant to include it as a factor for the depiction of invasive mammal's potential distribution (e.g. Cassinello *et al.*, 2006, López-Darias *et al.*, 2008). However, with time and propagule pressure, species can shift their niche after being introduced in similar climates as their native range. *C. stoebe,* one of the most important climatic niche shift observed in plants invaders (Broennimann *et al.*, 2007, Petitpierre *et al.*, 2012) has been historically introduced in similar conditions as its native range (Broennimann *et al.*, 2007). A recent study tested for niche conservatism between native and introduced ranges of the American grey squirrel, which is to our knowledge one of the rare test of niche conservatism for an invasive mammal species (Di Febbraro *et al.*, 2013, see Guisan *et al.*, 2014 for a review of niche comparisons between native and introduced ranges), supports niche shift between native and invaded range, where only half of its invaded distribution was predicted through ENM. With time, biological invasions may become less predictable, urging to take measures at early stage of invasion. Our results plead for a systematic inclusion of the native environmental niche in decision based-models used to optimize conservation strategies (Guisan *et al.*, 2013).

Limitation and perspective

Although our dataset is one of the rare including information about failed introduction, the prevalence between information about success and failure is highly biased, with about 10 fold more information about introduction success. However, the geographical distribution between successful and failed introduction does not show any particular bias (Fig. 3).

Species distribution modeling at the coarse resolution we used to depict the native niche can hardly catch important information about landuse, microhabitats, or micro topography of the native niche. Such factors can be more important for endothermic species like mammals than for reptiles, amphibians, invertebrates or plants where climate is assumed to be the main driver of species distribution at worldwide scale. Although environmental factors exist at such fine resolutions, finer information about species distribution is also required to build finer SDM. The resolution used in this study allows a fair compromise between comprehensiveness and precision about species distribution. Additionally, this scale is comparable to recent pre-border risk assessment used in conservation (Koop *et al.*, 2012).

CONCLUSION

Thanks to the availability of information about introduction's failure and success, we showed that species introductions are more successful in area with higher innerness with species' native niche. Although this signal for niche conservatism is only partial, the combination of information about native environmental niche matching with traits known for their roles in introduction success should provide more robust models of invasion success. Such information is elementary in a world where biodiversity is put at threat requiring efficient and optimized conservation planning based on prevention rather than eradication of invasive species.

REFERENCES

Allen, C.R., Nemec, K.T., Wardwell, D.A., Hoffman, J.D., Brust, M., Decker, K.L., Fogell, D., Hogue, J., Lotz, A., Miller, T., Pummill, M., Ramirez-Yanez, L.E. & Uden, D.R. (2013) Predictors of regional establishment success and spread of introduced non-indigenous vertebrates. *Global Ecology and Biogeography,* **22,** 889-899.

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

Araújo, M.B., Ferri-Yáñez, F., Bozinovic, F., Marquet, P.A., Valladares, F. & Chown, S.L. (2013) Heat freezes niche evolution. *Ecology Letters***, 16,** 1206-1219.

Araujo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution,* **22,** 42-47.

Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberon, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling,* **222,** 1810-1819.

Bininda-Emonds, O.R.P., Cardillo, M., Jones, K.E., MacPhee, R.D.E., Beck, R.M.D., Grenyer, R., Price, S.A., Vos, R.A., Gittleman, J.L. & Purvis, A. (2007) The delayed rise of present-day mammals. *Nature,* **446,** 507-512.

Blackburn, T.M. & Duncan, R.P. (2001) Determinants of establishment success in introduced birds. *Nature,* **414,** 195-197.

Blossey, B. & Notzold, R. (1995) Evolution of increased competitive ability in invasive nonindigenous plants: a hypothesis. *Journal of Ecology,* **83,** 887-889.

Bomford, M., Kraus, F., Barry, S.C. & Lawrence, E. (2009) Predicting establishment success for alien reptiles and amphibians: a role for climate matching. *Biological Invasions,* **11,** 713-724.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

Broennimann, O., Mráz, P., Petitpierre, B., Guisan, A. & Müller-Schärer, H. (2014) Contrasting spatio-temporal climatic niche dynamics during the eastern and western invasions of spotted knapweed in North America. *Journal of Biogeography,* **Early view**.

Broennimann, O., Treier, U.A., Muller-Scharer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters,* **10,** 701-709.

Cassinello, J., Acevedo, P. & Hortal, J. (2006) Prospects for population expansion of the exotic aoudad (Ammotragus lervia; Bovidae) in the Iberian Peninsula: clues from habitat suitability modelling. *Diversity and Distributions,* **12,** 666-678.

Chapple, D.G., Simmonds, S.M. & Wong, B.B.M. (2012) Can behavioral and personality traits influence the success of unintentional species introductions? *Trends in Ecology & Evolution,* **27,** 57-64.

Cooper, N., Freckleton, R.P. & Jetz, W. (2010) Phylogenetic conservatism of environmental niches in mammals. *Proceedings of the Royal Society B: Biological Sciences,* **278,** 2384-2391.

Cox, B.C. (2001) The biogeographic regions reconsidered. *Journal of Biogeography,* **28,** 511-523.

Di Febbraro, M., Lurz, P.W.W., Genovesi, P., Maiorano, L., Girardello, M. & Bertolino, S. (2013) The Use of Climatic Niches in Screening Procedures for Introduced Species to Evaluate Risk of Spread: A Case with the American Eastern Grey Squirrel. *Plos One,* **8,** e66559.

Dormann, C.F., Gruber, B., Winter, M. & Herrmann, D. (2010) Evolution of climate niches in European mammals? *Biology Letters,* **6,** 229-232.

Duncan, R.P., Blackburn, T.M. & Sol, D. (2003) The ecology of bird introductions. *Annual Review of Ecology Evolution and Systematics,* **34,** 71-98.

Fitzpatrick, M.C. & Hargrove, W.W. (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation,* **18,** 2255-2261.

Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography,* **16,** 24-33.

Forsyth, D.M. & Duncan, R.P. (2001) Propagule size and the relative success of exotic ungulate and bird introductions to New Zealand. *American Naturalist,* **157,** 583-595.

Forsyth, D.M., Duncan, R.P., Bomford, M. & Moore, G. (2004) Climatic suitability, life-history traits, introduction effort, and the establishment and spread of introduced mammals in Australia. *Conservation Biology,* **18,** 557-569.

Friedman, J.H. (2001) Greedy function approximation: A gradient boosting machine. *Annals of Statistics,* **29,** 1189-1232.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology,* **98,** 790-799.

Guisan, A., Petitpierre, B., Broennimann, O., Daehler, C. & Kueffer, C. (2014) Unifying niche shift studies: insights from biological invasions. *Trends in Ecology & Evolution,* **In Press**.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R.,

Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters,* **16,** 1424-1435.

Hadfield, J.D. (2010) MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *Journal of Statistical Software,* **33,** 1-22.

Hayes, K.R. & Barry, S.C. (2008) Are there any consistent predictors of invasion success? *Biological Invasions,* **10,** 483-506.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology,* **25,** 1965-1978.

Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2010) dismo: Species distribution modeling. pp. R package version 0.7-17. Available at: http://CRAN.R-project.org/package=dismo. Last accessed 8 August 2012.

Hulme, P.E. & Barrett, S.C.H. (2013) Integrating trait- and niche-based approaches to assess contemporary evolution in alien plant species. *Journal of Ecology,* **101,** 68-77.

Jimenez-Valverde, A., Peterson, A.T., Soberon, J., Overton, J.M., Aragon, P. & Lobo, J.M. (2011) Use of niche models in invasive species risk assessments. *Biological Invasions,* **13,** 2785-2797.

Kolar, C.S. & Lodge, D.M. (2001) Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution,* **16,** 199-204.

Kolar, C.S. & Lodge, D.M. (2002) Ecological predictions and risk assessment for alien fishes in North America. *Science,* **298,** 1233-1236.

Kolbe, J.J., VanMiddlesworth, P.S., Losin, N., Dappen, N. & Losos, J.B. (2012) Climatic niche shift predicts thermal trait response in one but not both introductions of the Puerto Rican lizard *Anolis cristatellus* to Miami, Florida, USA. *Ecology and Evolution,* **2,** 1503-1516.

Koop, A., Fowler, L., Newton, L. & Caton, B. (2012) Development and validation of a weed screening tool for the United States. *Biological Invasions,* **14,** 273-294.

Leung, B., Lodge, D.M., Finnoff, D., Shogren, J.F., Lewis, M.A. & Lamberti, G. (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **269,** 2407-2413.

Long, J. (2003) *Introduced Mammals of the World: Their History, Distribution and Influence.* CABI Publishing, Wallingord, UK.

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

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

Maron, J.L., Vila, M., Bommarco, R., Elmendorf, S. & Beardsley, P. (2004) Rapid evolution of an invasive plant. *Ecological Monographs,* **74,** 261-280.

McCullagh, P. & Nelder, J.A. (1989) *Generalized Linear Models. 2nd edition.* Chapman and Hall, London.

Medley, K.A. (2010) Niche shifts during the global invasion of the Asian tiger mosquito, Aedes albopictus Skuse (Culicidae), revealed by reciprocal distribution models. *Global Ecology and Biogeography,* **19,** 122-133.

Moles, A.T., Gruber, M.A.M. & Bonser, S.P. (2008) A new framework for predicting invasive plant species. *Journal of Ecology,* **96,** 13-17.

Myers, K. (1986) Introduced vertebrates in Australia with emphasis on the mammals. *Ecology of biological invasions: an Australian perspective* (eds R. H. Groves & J. J. Burdon), pp. 120:136. Australian Academy of Science, Canberra , Australia.

Normand, S., Treier, U.A., Randin, C., Vittoz, P., Guisan, A. & Svenning, J.-C. (2009) Importance of abiotic stress as a range-limit determinant for European plants: insights from species responses to climatic gradients. *Global Ecology and Biogeography,* **18,** 437-449.

Petersen, M.J. (2013) Evidence of a climatic niche shift following North American introductions of two crane flies (*Diptera*; genus *Tipula*). *Biological Invasions,* **15,** 885-897.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders. *Science,* **335,** 1344-1348.

Pheloung, P.C. (1996) CLIMATE: a system to predict the distribution of an organism based on climate preferences. Agriculture Western Australia Perth

Pheloung, P.C., Williams, P.A. & Halloy, S.R. (1999) A weed risk assessment

model for use as a biosecurity tool evaluating plant introductions. *Journal of Environmental Management,* **57,** 239–251.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* **190,** 231-259.

Prenter, J., MacNeil, C., Dick, J.T. & Dunn, A.M. (2004) Roles of parasites in animal invasions. *Trends in Ecology & Evolution,* **19,** 385-390.

Pyšek, P., Jarošík, V., Hulme, P.E., Kühn, I., Wild, J., Arianoutsou, M., Bacher, S., Chiron, F., Didžiulis, V., Essl, F., Genovesi, P., Gherardi, F., Hejda, M., Kark, S., Lambdon, P.W., Desprez-Loustau, M.-L., Nentwig, W., Pergl, J., Poboljšaj, K., Rabitsch, W., Roques, A., Roy, D.B., Shirley, S., Solarz, W., Vilà, M. & Winter, M. (2010) Disentangling the role of environmental and human pressures on biological invasions across Europe. *Proceedings of the National Academy of Sciences,* **107,** 12157-12162.

R (2005) *R : A language and environment for statistical computing. Version 2.5.1.* Development Core Team. R Foundation for Statistical Computing, Vienna, Austria.

Richardson, D.M., Pyšek, P., Rejmánek, M., Barbour, M.G., Panetta, F.D. & West, C.J. (2000) Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions,* **6,** 93-107.

Rödder, D. & Lötters, S. (2009) Niche shift versus niche conservatism? Climatic characteristics of the native and invasive ranges of the Mediterranean house gecko (*Hemidactylus turcicus*). *Global Ecology and Biogeography,* **18,** 674-687.

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S.D., Grosberg, R.K., HastingS, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I. & Rice, W.R. (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution,* **22,** 465-471.

Simberloff, D. (2009) We can eliminate invasions or live with them. Successful management projects. *Biological Invasions,* **11,** 149-157.

Simons, A.M. (2003) Invasive aliens and sampling bias. *Ecology Letters,* **6,** 278-280.

Sol, D., Bacher, S., Reader, S.M. & Lefebvre, L. (2008b) Brain size predicts the success of mammal species introduced into novel environments. *American Naturalist,* **172,** S63-S71.

Sutter, M. & Kawecki, T.J. (2009) Influence of learning on range expansion and adaptation to novel habitats. *Journal of Evolutionary Biology,* **22,** 2201-2214.

Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (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.

Vaclavik, T. & Meentemeyer, R.K. (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions,* **18,** 73-83.

Wolfe, L.M. (2002) Why alien invaders succeed: support for the escape from enemy hypothesis. *The American Naturalist,* **160,** 705-711.

Zomer, R.J., Trabucco, A., Bossio, D.A., van Straaten, O. & Verchot, L.V. (2008) Climate Change Mitigation: A Spatial Analysis of Global Land Suitability for Clean Development Mechanism Afforestation and Reforestation. *Agriculture, Ecosystems & Environment,* **126,** 67 - 80.

Zweig, M.H. & Campbell, G. (1993) Receiver-Operating Characteristic (Roc) Plots - a Fundamental Evaluation Tool in Clinical Medicine. *Clinical Chemistry,* **39,** 561-577.

SUPPLEMENTARY MATERIAL

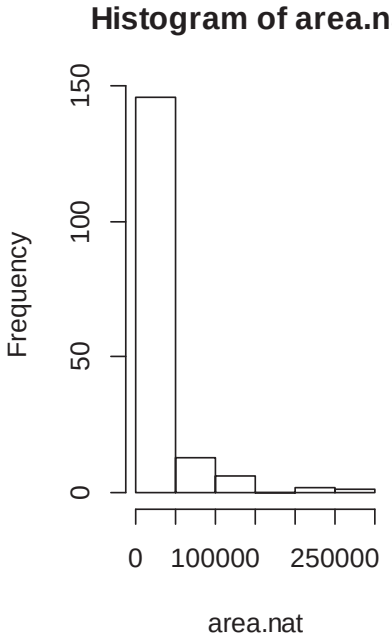


Figure S1: histogram of native area size

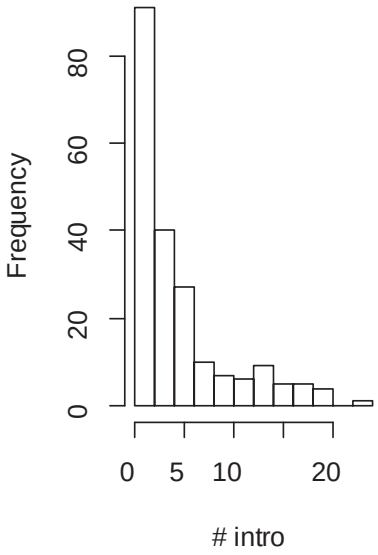
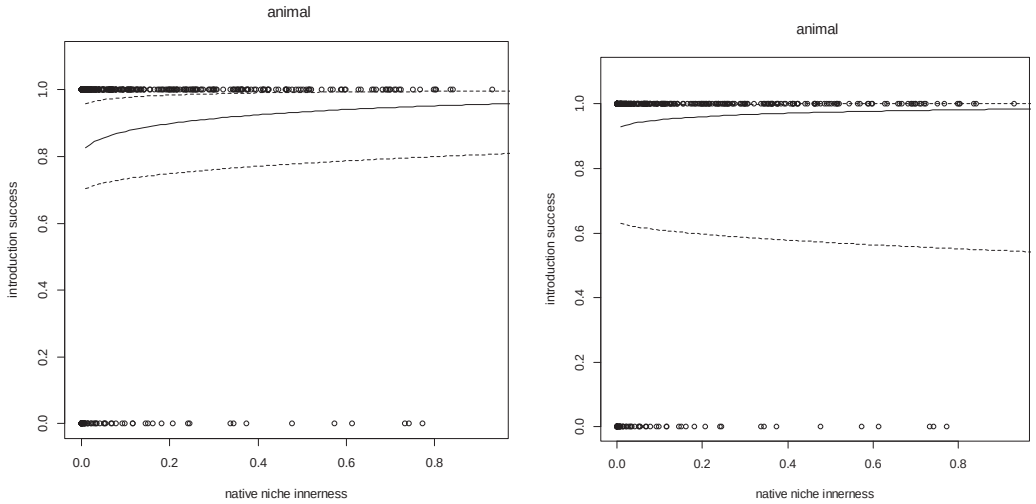


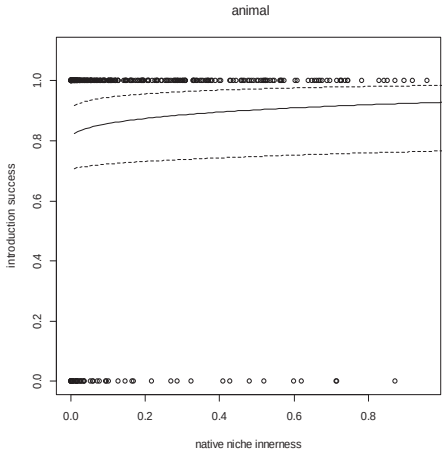
Figure S2: Number of introductions per species

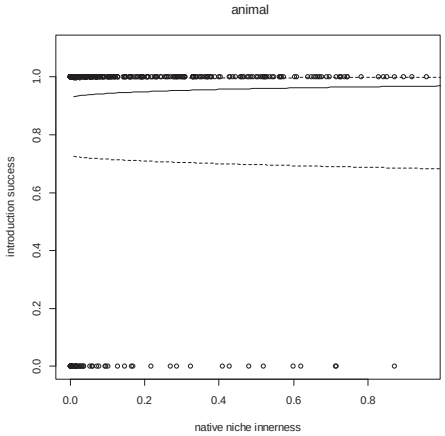



```
 DIC: 342.817

 G-structure:  ~animal

       post.mean  l-95% CI u-95% CI eff.samp
animal    0.4052 4.627e-07    3.866    10.23

 R-structure:  ~units

      post.mean l-95% CI u-95% CI eff.samp
units    0.1877 0.006194   0.9398    5.047

 Location effects: status ~ sqrt(dens)

            post.mean l-95% CI u-95% CI eff.samp pMCMC
(Intercept)    1.3809   0.8038   2.8675    93.82 0.016 *
sqrt(dens)     1.7829   0.6484   2.7609    10.49 0.004 **
---
```

```
 DIC: 357.3917

 G-structure:  ~animal

       post.mean l-95% CI u-95% CI eff.samp
animal     3.178  0.03639    17.39    6.668

 R-structure:  ~units

      post.mean l-95% CI u-95% CI eff.samp
units     2.688  0.06451    18.91    4.287

 Location effects: status ~ sqrt(suit)

            post.mean l-95% CI u-95% CI eff.samp pMCMC
(Intercept)    2.3948   0.5766   5.1414    10.52 0.002 **
sqrt(suit)     1.7862  -0.4151   4.8887    14.31 0.048 *
```

Figure S3: mixed effect model introduction success as a function of a) $INNER_{dens}$ and b) $INNER_{HS}$ (world calibration)






```
 DIC: 329.4308

 G-structure:  ~animal

       post.mean  l-95% CI u-95% CI eff.samp
animal    0.1862 2.079e-08   0.8213    7.664

 R-structure:  ~units

      post.mean  l-95% CI u-95% CI eff.samp
units    0.1148 0.0008833   0.7821    9.075

 Location effects: status ~ sqrt(dens)

            post.mean l-95% CI u-95% CI eff.samp  pMCMC
(Intercept)    1.4342   0.8498   2.2217     8.15 <0.001 ***
sqrt(dens)     1.1256   0.3353   1.9197    10.98  0.016 *
```

```
 DIC: 329.7755

 G-structure:  ~animal

       post.mean l-95% CI u-95% CI eff.samp
animal     2.698   0.2857    5.181    11.24

 R-structure:  ~units

      post.mean l-95% CI u-95% CI eff.samp
units    0.4162  0.01634    1.367    13.96

 Location effects: status ~ sqrt(suit)

             post.mean l-95% CI u-95% CI eff.samp pMCMC
(Intercept)     2.5077   0.9942   4.3129     7.96 0.002 **
sqrt(suit)      0.9273  -0.2312   2.0584    38.34 0.094 .
```

Figure S4: mixed effect model introduction success as a function of a) $INNER_{dens}$ and b) $INNER_{HS}$ (landuse calibration)

Table S1. List of the introductions including species name (sp.name), coordinates of the center of the introduction (intro.x and intro.y) and invasive status at each introduction site (0 = extinct, 1 = declining, 2 = established, 3 = expanding, 4 = eradicated). Status 0 and 1 were considered as failed introductions whereas 2 and 3 were considered as success. Status 4 was considered as NA.

sp.name	intro.x	intro.y	status	sp.name	intro.x	intro.y	status	sp.name	intro.x	intro.y	status
Acomys_cahirinus	24.92	35.25	3	Lasiorhinus_latifrons	135.58	-34.58	0	Oreamnos_americanus	-103.58	43.92	3
Acomys_cahirinus	33.25	35.08	3	Lasiorhinus_latifrons	137.25	-35.75	4	Oreamnos_americanus	-107.67	43.50	3
Alces_alces	-146.08	63.08	3	Eulemur_fulvus	45.08	-12.75	2	Oreamnos_americanus	-106.42	38.92	3
Alces_alces	-134.92	58.75	3	Eulemur_fulvus	49.75	-15.42	3	Oreamnos_americanus	-118.42	33.42	3
Alces_alces	-54.42	49.25	5	Eulemur_fulvus	46.25	-25.00	2	Ornithorhynchus_anatinus	136.83	-35.92	2
Alces_alces	-57.08	49.08	3	Eulemur_mongoz	43.33	-11.75	2	Ornithorhynchus_anatinus	116.92	-32.17	2
Alces_alces	-73.92	44.08	4	Eulemur_mongoz	44.42	-12.25	2	Oryx_gazella	-106.25	32.75	3
Alces_alces	-153.42	57.58	4	Lepus_americanus	-152.75	58.08	0	Ovibos_moschatus	-71.58	78.58	3
Alces_alces	36.08	51.75	0	Lepus_americanus	-152.75	58.25	2	Ovibos_moschatus	-69.58	76.42	3
Alces_alces	-63.00	49.42	3	Lepus_americanus	-134.42	57.58	2	Ovibos_moschatus	-147.25	70.08	3
Alces_alces	-60.92	46.25	3	Lepus_americanus	-153.42	57.58	2	Ovibos_moschatus	-143.75	70.08	3
Alopex_lagopus	54.25	60.08	3	Lepus_americanus	-63.00	49.42	2	Ovibos_moschatus	-50.75	67.08	3
Alopex_lagopus	-166.33	60.08	3	Lepus_americanus	-74.75	40.25	2	Ovibos_moschatus	5.75	62.25	4
Alopex_lagopus	-169.58	56.58	3	Lepus_americanus	-70.58	41.42	2	Ovibos_moschatus	-178.92	71.25	3
Alopex_lagopus	166.25	55.00	3	Lepus_americanus	-70.08	41.25	2	Ovibos_moschatus	-166.33	60.08	3
Ammotragus_lervia	-105.92	37.08	2	Lepus_americanus	-131.08	55.75	2	Ovis_ammon	10.42	53.42	2
Ammotragus_lervia	-103.58	36.25	3	Lepus_arcticus	-55.42	47.75	0	Ovis_ammon	-6.25	53.25	4
Ammotragus_lervia	-104.42	35.92	3	Lepus_arcticus	-55.58	50.83	2	Ovis_ammon	-3.58	50.42	0
Ammotragus_lervia	-107.58	35.25	2	Lepus_californicus	-0.58	52.08	0	Ovis_ammon	14.25	49.75	3
Ammotragus_lervia	-105.42	33.42	2	Lepus_californicus	-74.75	40.25	3	Ovis_ammon	1.58	47.58	3
Ammotragus_lervia	-105.25	33.42	2	Lepus_californicus	-70.58	41.42	3	Ovis_ammon	-97.92	27.58	3
Ammotragus_lervia	-100.75	27.08	3	Lepus_californicus	-70.08	41.25	3	Ovis_ammon	8.58	47.42	4
Ammotragus_lervia	-100.92	26.92	3	Lepus_californicus	-84.92	37.92	3	Ovis_ammon	9.08	42.25	3
Ammotragus_lervia	-106.08	36.42	3	Lepus_californicus	-59.42	-51.75	2	Ovis_ammon	-70.25	-38.58	2
Ammotragus_lervia	-106.42	35.25	3	Lepus_europaeus	113.42	52.08	2	Ovis_ammon	-73.17	-40.58	2
Ammotragus_lervia	-105.58	33.00	3	Lepus_europaeus	-89.25	48.42	3	Ovis_canadensis	-119.58	50.92	3
Ammotragus_lervia	-104.75	31.75	2	Lepus_europaeus	-73.75	41.75	3	Ovis_canadensis	-121.42	50.42	3
Ammotragus_lervia	-17.75	28.75	2	Lepus_europaeus	-73.75	41.08	3	Ovis_canadensis	-153.42	57.58	4
Ammotragus_lervia	-100.92	22.08	3	Lepus_europaeus	-75.42	40.58	3	Ovis_canadensis	-123.58	51.25	3
Antilocapra_americana	-76.75	39.25	2	Lepus_europaeus	-61.42	-32.75	3	Ovis_canadensis	-114.25	47.92	3
Atlantoxerus_getulus	-13.92	28.42	2	Lepus_europaeus	-2.92	59.08	2	Ovis_canadensis	-114.08	36.67	3
Bandicota_indica	120.92	23.75	2	Lepus_europaeus	-6.67	58.08	2	Ovis_canadensis	-115.58	34.25	3

Bandicota_indica	100.58	6.00	3	Lepus_europaeus	-5.92	56.58	4	Paguma_larvata	133.42	33.75	3
Bettongia_gaimardi	148.08	-42.58	2	Lepus_europaeus	-4.58	54.25	2	Paguma_larvata	130.08	32.42	3
Boselaphus_tragocamelus	12.50	41.92	4	Lepus_europaeus	-70.58	41.42	3	Pan_troglodytes	5.92	51.92	2
Callithrix_jacchus	-43.25	-22.92	3	Lepus_europaeus	-70.08	41.25	3	Pan_troglodytes	-81.17	31.83	2
Callosciurus_erythraeus	139.42	34.75	2	Lepus_europaeus	-61.67	16.17	3	Pan_troglodytes	-80.25	26.75	2
Canis_aureus	144.92	-37.75	4	Lepus_nigricollis	106.75	-6.25	2	Pan_troglodytes	32.58	-0.08	2
Capra_ibex	-5.42	57.25	5	Lepus_nigricollis	55.58	-21.08	3	Pan_troglodytes	31.83	-2.33	2
Capra_ibex	-106.58	35.08	2	Lepus_timidus	-5.92	57.25	0	Paradoxurus_hermaphroditus	127.92	1.08	3
Capra_ibex	34.25	45.25	3	Lepus_timidus	-4.08	55.42	3	Paradoxurus_hermaphroditus	127.42	-0.58	3
Capreolus_capreolus	-6.25	55.75	3	Lepus_timidus	-1.92	53.58	3	Paradoxurus_hermaphroditus	124.83	-1.92	3
Capreolus_capreolus	-8.58	54.08	4	Lepus_timidus	-1.58	53.58	3	Paradoxurus_hermaphroditus	129.42	-3.25	3
Capreolus_capreolus	-9.42	53.92	4	Lepus_timidus	104.25	52.25	2	Paradoxurus_hermaphroditus	132.75	-5.58	3
Capreolus_capreolus	-73.75	41.75	2	Lepus_timidus	6.58	50.92	0	Paradoxurus_hermaphroditus	134.42	-6.25	3
Capreolus_capreolus	-120.08	33.92	2	Lepus_timidus	8.08	50.92	0	Perameles_gunnii	148.08	-42.58	2
Castor_canadensis	27.58	62.42	3	Lepus_timidus	6.42	50.58	0	Petaurus_breviceps	127.42	0.75	3
Castor_canadensis	26.25	50.58	0	Lepus_timidus	-6.92	62.08	3	Petaurus_breviceps	143.92	-35.58	3
Castor_canadensis	3.08	47.58	3	Lepus_timidus	-6.25	57.25	3	Petaurus_breviceps	146.58	-41.92	3
Castor_canadensis	-132.25	53.58	0	Lepus_timidus	-5.83	56.00	3	Petrogale_penicillata	175.42	-36.25	4
Castor_canadensis	-63.00	49.42	3	Lepus_timidus	-6.25	55.75	0	Phascolarctos_cinereus	140.58	-34.25	4
Castor_canadensis	-68.58	-54.08	3	Lepus_timidus	-5.25	55.58	3	Phascolarctos_cinereus	138.75	-35.25	3
Cavia_aperea	-90.42	-0.58	2	Lepus_timidus	137.50	55.00	2	Phascolarctos_cinereus	140.25	-36.92	2
Cebus_albifrons	-61.25	10.42	2	Lepus_timidus	-3.08	54.50	0	Phascolarctos_cinereus	148.92	-20.92	2
Cebus_apella	-63.92	10.92	2	Loxodonta_africana	30.58	-1.58	3	Phascolarctos_cinereus	139.92	-34.08	3
Chlorocebus_aethiops	-23.67	15.08	2	Lycaon_pictus	31.33	-27.50	3	Phascolarctos_cinereus	137.25	-35.75	3
Chlorocebus_aethiops	-59.58	13.17	2	Macaca_arctoides	-95.08	18.42	2	Phascolarctos_cinereus	144.25	-37.92	2
Cercopithecus_mona	-61.75	12.08	2	Macaca_fascicularis	-80.42	25.58	3	Phascolarctos_cinereus	144.92	-40.42	4
Axis_axis	-122.75	38.08	3	Macaca_fascicularis	114.25	22.25	3	Potamochoerus_porcus	45.08	-12.75	3
Axis_axis	146.25	-20.08	3	Macaca_fascicularis	121.00	-8.75	3	Potorous_tridactylus	148.08	-42.58	2
Axis_axis	142.25	-36.75	3	Macaca_fascicularis	57.58	-20.25	3	Potorous_tridactylus	167.83	-47.08	0
Axis_axis	-0.83	51.75	4	Macaca_fuscata	-99.42	27.92	2	Procyon_lotor	-3.08	51.92	2
Axis_axis	37.58	48.08	3	Macaca_mulatta	13.42	52.58	4	Procyon_lotor	8.92	51.08	3
Axis_axis	33.58	46.75	3	Macaca_mulatta	7.08	50.75	4	Procyon_lotor	-0.08	50.92	2
Axis_axis	74.00	31.08	3	Macaca_mulatta	115.75	40.42	2	Procyon_lotor	6.75	50.42	3
Axis_axis	-81.25	29.08	3	Macaca_mulatta	-80.75	32.25	4	Procyon_lotor	176.25	-38.08	0
Axis_axis	92.75	12.42	3	Macaca_mulatta	-82.08	29.25	2	Procyon_lotor	-153.42	57.58	3
Axis_axis	145.17	-5.17	2	Macaca_mulatta	-82.75	21.75	0	Procyon_lotor	-132.42	53.75	3
Axis_axis	-57.92	-34.25	3	Macaca_mulatta	-66.00	18.42	2	Procyon_lotor	0.92	52.58	2

Axis_axis	145.08	-37.08	3	Macaca_nemestrina	92.92	12.25	2	Procyon_lotor	8.25	46.75	3
Cervus_elaphus	27.92	-26.58	3	Macaca_nemestrina	100.25	5.42	2	Procyon_lotor	46.75	42.92	3
Cervus_elaphus	152.42	-27.08	3	Macaca_nemestrina	103.83	1.42	2	Procyon_lotor	43.25	43.42	3
Cervus_elaphus	152.42	-28.08	3	Macropus_agilis	152.58	-3.92	2	Procyon_lotor	73.58	41.25	3
Cervus_elaphus	152.08	-28.25	3	Macropus_agilis	151.08	-8.58	2	Procyon_lotor	-80.92	34.08	5
Cervus_elaphus	115.92	-32.58	5	Macropus_agilis	150.25	-9.33	2	Procyon_lotor	-77.67	24.50	3
Cervus_elaphus	115.92	-32.75	5	Macropus_agilis	150.67	-9.50	2	Procyon_lotor	-61.67	16.17	2
Cervus_elaphus	140.75	-37.42	4	Macropus_agilis	149.92	-9.75	2	Procyon_lotor	-59.58	13.17	2
Cervus_elaphus	143.75	-37.92	2	Macropus_fuliginosus	115.92	-31.92	2	Pseudois_nayaur	170.08	-43.58	4
Cervus_elaphus	145.58	-37.92	4	Macropus_giganteus	148.92	-20.25	5	Pseudocheirus_occidentalis	151.08	-33.75	2
Cervus_elaphus	143.58	-38.75	5	Macropus_giganteus	148.08	-42.58	3	Pseudocheirus_occidentalis	137.25	-35.75	2
Cervus_elaphus	8.75	3.50	2	Macropus_giganteus	169.25	-45.75	0	Pseudocheirus_occidentalis	146.58	-41.92	5
Cervus_elaphus	146.25	-18.25	4	Macropus_giganteus	168.42	-46.58	0	Pseudocheirus_occidentalis	167.83	-47.08	2
Odocoileus_hemionus	-0.58	52.08	0	Macropus_robustus	148.92	-20.17	4	Rangifer_tarandus	-14.92	65.75	3
Odocoileus_hemionus	-153.42	57.58	3	Macropus_robustus	137.25	-35.75	4	Rangifer_tarandus	-153.42	57.58	2
Odocoileus_hemionus	-132.42	53.75	3	Macropus_rufogriseus	-4.58	56.08	0	Rangifer_tarandus	148.75	45.42	5
Odocoileus_hemionus	-131.92	52.92	3	Macropus_rufogriseus	14.08	51.58	4	Rangifer_tarandus	-67.75	-54.42	4
Odocoileus_hemionus	-74.42	44.08	4	Macropus_rufogriseus	-0.25	51.08	0	Rangifer_tarandus	-36.67	-54.33	3
Odocoileus_hemionus	-96.92	28.17	4	Macropus_rufogriseus	-0.08	51.08	4	Rattus_argentiventer	120.42	-8.75	3
Odocoileus_hemionus	175.08	-37.25	4	Macropus_rufogriseus	6.92	50.75	4	Rattus_nitidus	121.08	16.42	3
Odocoileus_hemionus	168.58	-45.08	4	Macropus_rufogriseus	8.42	49.75	4	Rattus_nitidus	134.58	7.50	3
Rusa_timorensis	101.42	-0.58	4	Macropus_rufogriseus	148.08	-42.58	3	Rattus_nitidus	129.42	-3.25	3
Rusa_timorensis	114.92	-3.25	1	Macropus_rufogriseus	169.25	-44.58	3	Rattus_nitidus	137.08	-4.08	3
Rusa_timorensis	152.08	-4.25	3	Macropus_rufogriseus	171.08	-44.75	3	Rattus_praetor	152.58	-3.92	2
Rusa_timorensis	134.25	-5.75	3	Macropus_rufogriseus	-5.17	55.92	4	Rattus_praetor	155.25	-6.25	2
Rusa_timorensis	142.25	-9.25	3	Macrotis_lagotis	136.92	-30.58	3	Rattus_praetor	157.00	-7.08	2
Rusa_timorensis	149.58	-32.42	3	Marmota_bobak	55.92	58.08	0	Rattus_praetor	160.25	-9.58	2
Rusa_timorensis	115.08	-34.25	0	Marmota_bobak	39.58	49.25	0	Rattus_tanezumi	145.75	15.25	3
Rusa_timorensis	146.75	-38.25	0	Marmota_bobak	46.92	42.42	2	Rattus_tanezumi	144.83	13.50	3
Rusa_timorensis	124.83	-1.92	3	Martes_americana	-152.75	58.25	3	Rattus_tanezumi	158.25	6.92	3
Rusa_timorensis	126.67	-3.42	3	Martes_americana	-135.75	57.92	3	Rattus_tanezumi	127.42	0.75	3
Rusa_timorensis	128.17	-3.58	3	Martes_americana	-153.42	57.58	4	Rattus_tanezumi	127.67	-1.58	3
Rusa_timorensis	125.00	-9.25	3	Martes_americana	-135.08	57.08	3	Rattus_tanezumi	125.92	-2.08	3
Rusa_timorensis	44.42	-12.25	2	Martes_americana	-132.75	55.58	3	Rhinolophus_ferrumequinum	-0.08	51.58	2
Rusa_unicolor	132.08	-11.42	3	Martes_americana	-88.58	45.42	2	Rhinolophus_ferrumequinum	-6.92	54.08	2
Rusa_unicolor	-120.33	35.25	5	Martes_martes	-4.42	56.92	0	Rhizomys_sumatrensis	103.83	1.42	2
Rusa_unicolor	-85.17	29.75	3	Martes_melampus	138.42	38.08	2	Rupicapra_rupicapra	8.25	48.25	2

Rusa_unicolor	144.83	13.50	4	Martes_pennanti	-63.00	49.42	3	Saimiri_sciureus	-69.92	-4.08	1
Rusa_unicolor	158.25	6.92	2	Meles_meles	-2.75	54.75	2	Saimiri_sciureus	-83.17	9.75	2
Rusa_unicolor	28.58	-20.58	2	Meles_meles	-4.42	58.42	2	Saimiri_sciureus	-77.75	8.58	1
Rusa_unicolor	142.42	-37.25	3	Meles_meles	-5.92	55.92	0	Saimiri_sciureus	-82.42	8.25	1
Rusa_unicolor	145.08	-37.08	3	Meles_meles	-4.75	54.92	2	Sciurus_aberti	-110.58	32.25	2
Rusa_unicolor	143.58	-37.42	3	Meles_meles	-2.92	54.58	2	Sciurus_carolinensis	-3.42	55.92	3
Rusa_unicolor	176.25	-38.08	3	Meles_meles	-1.42	50.75	4	Sciurus_carolinensis	-3.25	55.92	3
Rusa_unicolor	146.58	-38.75	3	Mephitis_mephitis	-125.58	49.75	2	Sciurus_carolinensis	-7.92	53.75	3
Rusa_unicolor	175.25	-40.25	3	Mephitis_mephitis	36.42	49.58	2	Sciurus_carolinensis	-7.75	53.58	3
Odocoileus_virginianus	-0.58	52.08	4	Mephitis_mephitis	-63.08	46.42	3	Sciurus_carolinensis	-2.08	53.42	3
Odocoileus_virginianus	14.25	49.75	3	Mephitis_mephitis	-63.75	45.08	2	Sciurus_carolinensis	-0.58	51.58	3
Odocoileus_virginianus	-61.42	15.42	2	Mephitis_mephitis	72.25	40.25	2	Sciurus_carolinensis	-0.42	51.42	0
Odocoileus_virginianus	168.42	-45.08	2	Meriones_unguiculatus	-0.83	53.67	0	Sciurus_carolinensis	-0.25	51.42	3
Odocoileus_virginianus	167.83	-47.08	3	Meriones_unguiculatus	-1.08	53.58	0	Sciurus_carolinensis	-123.08	49.25	3
Myodes_glareolus	-9.17	52.08	3	Meriones_unguiculatus	-1.25	50.75	0	Sciurus_carolinensis	-111.25	47.58	3
Myodes_glareolus	-2.08	49.25	2	Miopithecus_talapoin	-15.58	28.25	5	Sciurus_carolinensis	-102.92	47.42	3
Myodes_rutilus	166.25	55.08	3	Miopithecus_talapoin	8.75	3.50	5	Sciurus_carolinensis	-98.75	46.92	3
Connochaetes_taurinus	33.92	46.42	3	Moschus_moschiferus	-0.58	51.92	4	Sciurus_carolinensis	-98.08	46.92	3
Connochaetes_taurinus	9.75	-0.58	3	Mungos_mungo	39.42	-6.00	2	Sciurus_carolinensis	-100.75	46.75	3
Suncus_murinus	144.83	13.50	3	Muntiacus_muntjak	116.25	-8.58	0	Sciurus_carolinensis	7.58	44.92	3
Suncus_murinus	39.25	-6.08	3	Muntiacus_reevesi	-0.58	52.08	0	Sciurus_carolinensis	18.42	-33.92	3
Suncus_murinus	43.42	-11.75	0	Muntiacus_reevesi	-0.58	51.92	0	Sciurus_carolinensis	143.92	-37.58	0
Suncus_murinus	55.58	-21.08	0	Muntiacus_reevesi	-0.08	51.75	0	Sciurus_carolinensis	-125.58	49.75	3
Crocidura_suaveolens	-2.08	49.25	2	Muntiacus_reevesi	-0.92	52.33	0	Sciurus_niger	-102.92	47.42	3
Dasyprocta_leporina	-61.75	12.08	3	Muscardinus_avellanarius	0.92	52.58	1	Sciurus_niger	-98.75	46.92	3
Dasypus_novemcinctus	-80.92	29.08	3	Mustela_erminea	-1.25	60.42	3	Sciurus_niger	-76.75	43.08	3
Dasypus_novemcinctus	-80.25	25.92	3	Mustela_erminea	-2.92	59.08	2	Sciurus_niger	-73.92	40.75	3
Dasypus_novemcinctus	-80.92	34.08	3	Mustela_erminea	5.42	53.42	3	Sciurus_niger	-119.75	36.75	3
Dasypus_novemcinctus	-80.75	28.25	3	Mustela_nivalis	5.42	53.42	0	Sciurus_niger	-119.08	34.58	3
Dasypus_novemcinctus	144.83	13.50	3	Mustela_nivalis	2.92	39.58	2	Sciurus_niger	-118.42	34.08	3
Daubentonia_madagascariensis	49.75	-15.42	2	Mustela_nivalis	-27.92	38.58	2	Sciurus_vulgaris	37.58	55.75	2
Dendrolagus_matschiei	147.92	-5.58	2	Mustela_nivalis	14.42	35.92	2	Sciurus_vulgaris	41.92	45.08	2
Didelphis_marsupialis	-122.25	48.58	3	Mustela_nivalis	6.67	0.25	2	Sciurus_vulgaris	14.00	54.00	2
Didelphis_marsupialis	-121.92	37.25	3	Mustela_putorius	147.08	-41.42	0	Sciurus_vulgaris	34.25	45.25	2
Didelphis_marsupialis	-111.58	33.42	3	Mustela_putorius	140.42	38.75	0	Sciurus_vulgaris	43.58	42.92	3
Didelphis_marsupialis	-110.92	32.25	3	Mustela_putorius	9.08	40.08	3	Suncus_etruscus	100.50	13.75	2
Didelphis_marsupialis	-6.08	56.42	2	Mustela_putorius	-77.33	18.17	0	Spermophilus_beecheyi	170.42	-45.92	0

Didelphis_marsupialis	-122.42	48.25	3	Mustela_putorius	171.92	-43.67	3	Spermophilus_major	43.00	43.08	3
Didelphis_marsupialis	-78.58	26.75	2	Mustela_sibirica	142.58	50.42	3	Spermophilus_parryii	-153.42	57.58	2
Didelphis_marsupialis	-61.42	15.42	2	Mustela_sibirica	142.58	43.25	3	Spermophilus_parryii	-169.58	56.58	2
Didelphis_marsupialis	-60.92	14.00	2	Mustela_sibirica	77.42	42.92	3	Spermophilus_parryii	-166.75	53.75	2
Didelphis_marsupialis	-61.25	13.25	2	Mustela_sibirica	77.92	42.75	3	Spilocuscus_maculatus	149.58	-1.42	2
Didelphis_marsupialis	-61.75	12.08	2	Mustela_sibirica	135.25	34.75	3	Spilocuscus_maculatus	149.58	-1.58	2
Dipodomys_ordii	-81.58	42.08	2	Mustela_sibirica	133.42	33.75	3	Spilocuscus_maculatus	152.58	-3.92	2
Pseudalopex_griseus	-69.42	-52.92	3	Mustela_sibirica	129.42	28.25	0	Spilocuscus_maculatus	128.17	-3.58	2
Pseudalopex_griseus	-68.58	-52.92	3	Mustela_sibirica	124.25	24.42	3	Spilocuscus_maculatus	132.75	-5.58	2
Pseudalopex_griseus	-61.00	-51.92	3	Neovison_vison	19.25	63.92	3	Sus_scrofa	40.75	56.08	3
Makalata_didelphoides	-60.92	14.75	0	Neovison_vison	13.42	60.92	3	Sus_scrofa	-132.75	55.58	3
Echymipera_kalubu	147.00	-2.08	5	Neovison_vison	12.08	58.25	3	Sus_scrofa	32.92	55.08	3
Equus_quagga	32.33	-26.92	3	Neovison_vison	-55.42	47.58	3	Sus_scrofa	36.08	51.75	3
Erinaceus_europaeus	-0.92	60.58	4	Neovison_vison	-1.33	60.50	4	Sus_scrofa	0.58	51.25	4
Erinaceus_europaeus	174.08	-39.08	3	Neovison_vison	20.08	60.25	3	Sus_scrofa	-83.42	35.25	3
Erinaceus_europaeus	175.75	-39.42	3	Neovison_vison	-147.50	60.00	2	Sus_scrofa	-66.42	18.25	3
Erinaceus_europaeus	173.08	-41.25	3	Neovison_vison	-6.67	58.08	3	Sus_scrofa	-89.50	-0.83	3
Erinaceus_europaeus	171.08	-43.75	5	Neovison_vison	-153.42	57.58	0	Sus_scrofa	130.50	-0.92	3
Erinaceus_europaeus	169.25	-45.08	3	Neovison_vison	16.58	56.75	3	Sus_scrofa	126.67	-3.42	3
Erinaceus_europaeus	169.25	-45.25	3	Neovison_vison	-5.25	55.58	3	Sus_scrofa	150.58	-5.58	3
Erinaceus_europaeus	170.58	-45.75	5	Neovison_vison	-7.42	54.58	3	Sus_scrofa	134.42	-6.25	3
Erinaceus_europaeus	-1.00	60.58	2	Neovison_vison	-6.42	53.42	3	Sus_scrofa	117.75	-8.58	3
Erinaceus_europaeus	-3.08	59.08	2	Neovison_vison	148.75	45.42	3	Sus_scrofa	161.75	-10.58	3
Erinaceus_europaeus	-5.92	56.58	2	Neovison_vison	-63.00	49.42	0	Sus_scrofa	137.25	-35.75	3
Erinaceus_europaeus	-5.17	55.92	2	Neovison_vison	142.58	43.25	2	Sylvilagus_floridanus	-123.25	44.58	3
Erinaceus_europaeus	167.83	-47.08	3	Neovison_vison	-72.25	-41.08	0	Sylvilagus_floridanus	-76.92	38.92	3
Euphractus_sexcinctus	-72.08	-36.75	2	Myocastor_coypus	0.25	52.58	3	Sylvilagus_floridanus	-125.58	49.75	3
Lynx_lynx	11.58	48.17	2	Myocastor_coypus	0.25	52.25	3	Sylvilagus_floridanus	-70.58	41.42	3
Funambulus_pennantii	151.25	-33.92	0	Myocastor_coypus	-0.25	51.08	3	Sylvilagus_floridanus	-70.08	41.25	3
Gazella_subgutturosa	-106.58	35.08	2	Myocastor_coypus	-122.75	49.08	3	Sylvilagus_floridanus	-76.75	39.25	3
Gazella_subgutturosa	50.58	26.00	3	Myocastor_coypus	-94.58	48.75	3	Sylvilagus_floridanus	-69.08	12.25	3
Genetta_genetta	2.92	39.58	3	Myocastor_coypus	-74.25	45.58	3	Tachyglossus_aculeatus	148.25	-40.25	5
Genetta_genetta	1.42	39.00	3	Myocastor_coypus	-118.42	34.58	0	Tachyglossus_aculeatus	148.08	-42.58	2
Giraffa_camelopardalis	31.83	-2.33	2	Myocastor_coypus	35.58	32.42	3	Talpa_europaea	78.58	55.08	2
Giraffa_camelopardalis	31.92	-28.25	2	Myocastor_coypus	-95.42	31.58	3	Talpa_europaea	-5.92	56.58	2
Glis_glis	-0.58	51.75	3	Myocastor_coypus	11.42	52.08	3	Talpa_europaea	-6.08	56.42	0
Glis_glis	-0.25	51.58	3	Myocastor_coypus	0.92	52.58	3	Talpa_europaea	-5.17	55.92	2

Hemitragus_jemlahicus	-4.42	54.92	4	Myocastor_coypus	0.92	52.08	3	Talpa_europaea	37.58	48.08	2
Hemitragus_jemlahicus	170.08	-43.58	3	Myocastor_coypus	-122.25	47.67	3	Tamias_sibiricus	6.08	46.25	2
Hemitragus_jemlahicus	-78.33	44.25	0	Myocastor_coypus	7.25	47.25	5	Tamias_sibiricus	140.25	35.50	3
Hemitragus_jemlahicus	18.42	-33.92	2	Myocastor_coypus	-123.75	45.42	3	Tamias_striatus	-80.50	41.58	2
Herpestes_javanicus	-84.75	38.08	4	Myocastor_coypus	-82.33	38.92	3	Tamiasciurus_hudsonicus	-56.08	50.92	3
Herpestes_javanicus	-121.08	36.58	4	Myocastor_coypus	-121.00	37.58	0	Tamiasciurus_hudsonicus	-152.75	58.25	3
Herpestes_javanicus	-80.25	25.75	4	Myocastor_coypus	8.33	49.08	3	Tamiasciurus_hudsonicus	-135.75	57.92	3
Herpestes_javanicus	-71.25	18.92	2	Myocastor_coypus	-82.42	28.00	3	Tamiasciurus_hudsonicus	-153.42	57.58	3
Herpestes_javanicus	-77.33	18.17	3	Myrmecobius_fasciatus	118.08	-34.42	3	Tamiasciurus_hudsonicus	-135.08	57.17	3
Herpestes_javanicus	-66.50	18.25	2	Nasua_nasua	-86.25	39.92	3	Tayassu_pecari	-83.75	22.42	3
Herpestes_javanicus	-64.75	17.75	2	Nesotragus_moschatus	31.92	-28.25	4	Pecari_tajacu	-4.58	56.08	4
Herpestes_javanicus	-61.75	17.33	0	Nyctereutes_procyonoides	31.58	59.92	3	Pecari_tajacu	-86.83	20.50	5
Herpestes_javanicus	-61.58	16.25	2	Nyctereutes_procyonoides	43.75	58.58	3	Pecari_tajacu	9.42	-0.58	2
Herpestes_javanicus	-61.42	15.42	0	Nyctereutes_procyonoides	32.42	58.25	3	Tenrec_ecaudatus	45.08	-12.75	2
Herpestes_javanicus	-60.92	14.75	2	Nyctereutes_procyonoides	29.25	57.25	3	Tenrec_ecaudatus	57.58	-20.25	2
Herpestes_javanicus	-60.92	14.00	3	Nyctereutes_procyonoides	39.08	57.92	3	Tenrec_ecaudatus	55.58	-21.08	2
Herpestes_javanicus	-61.25	13.25	3	Nyctereutes_procyonoides	52.75	57.08	3	Thylogale_browni	152.58	-3.92	2
Herpestes_javanicus	-59.58	13.17	3	Nyctereutes_procyonoides	47.08	55.58	3	Thylogale_browni	152.58	-3.08	4
Herpestes_javanicus	-61.75	12.08	2	Nyctereutes_procyonoides	32.92	54.92	3	Thylogale_browni	150.58	-5.58	2
Herpestes_javanicus	-61.25	10.42	3	Nyctereutes_procyonoides	40.58	54.25	3	Thylogale_browni	154.58	-5.17	4
Herpestes_javanicus	-53.25	3.92	3	Nyctereutes_procyonoides	37.58	53.92	3	Thylogale_browni	147.92	-5.58	2
Herpestes_javanicus	128.17	-3.58	2	Nyctereutes_procyonoides	33.58	52.92	3	Trachypithecus_auratus	116.25	-8.58	3
Herpestes_javanicus	39.75	-7.92	2	Nyctereutes_procyonoides	44.58	53.25	3	Tragelaphus_strepsiceros	-106.58	35.08	2
Herpestes_javanicus	179.33	-16.58	3	Nyctereutes_procyonoides	23.92	52.75	3	Trichosurus_vulpecula	176.08	-37.75	2
Herpestes_javanicus	178.00	-17.92	3	Nyctereutes_procyonoides	47.25	47.25	3	Trichosurus_vulpecula	148.92	-20.92	2
Herpestes_javanicus	57.58	-20.25	2	Nyctereutes_procyonoides	43.42	43.42	3	Trichosurus_vulpecula	173.75	-41.08	2
Herpestes_javanicus	143.08	-35.08	0	Nycticebus_coucang	124.92	7.75	2	Trichosurus_vulpecula	148.08	-42.58	2
Herpestes_edwardsii	13.08	41.25	3	Ondatra_zibethicus	23.08	59.92	3	Trichosurus_vulpecula	168.42	-44.92	2
Herpestes_edwardsii	127.92	26.50	3	Ondatra_zibethicus	-5.42	56.42	4	Trichosurus_vulpecula	166.08	-50.75	0
Herpestes_edwardsii	101.25	4.75	3	Ondatra_zibethicus	90.58	53.08	5	Varecia_variegata	49.75	-15.42	2
Herpestes_edwardsii	100.58	5.25	3	Ondatra_zibethicus	-2.75	52.75	4	Viverra_tangalunga	127.92	1.08	2
Herpestes_edwardsii	101.75	3.25	3	Ondatra_zibethicus	14.25	49.75	3	Viverra_tangalunga	127.58	-0.58	2
Herpestes_edwardsii	102.42	2.42	3	Ondatra_zibethicus	-123.25	48.75	3	Viverra_tangalunga	126.67	-3.42	3
Herpestes_edwardsii	57.58	-20.25	3	Ondatra_zibethicus	126.75	45.75	3	Viverra_zibetha	92.75	12.42	3
Herpestes_edwardsii	168.08	-45.58	0	Ondatra_zibethicus	27.08	44.08	3	Viverricula_indica	53.92	12.58	2
Herpestes_edwardsii	143.08	-35.08	0	Ondatra_zibethicus	35.67	65.08	3	Viverricula_indica	39.42	-6.00	2
Herpestes_ichneumon	16.92	42.92	2	Ondatra_zibethicus	164.25	58.92	3	Viverricula_indica	115.25	-8.25	2

Herpestes_ichneumon	17.58	42.75	0	Ondatra_zibethicus	-153.42	57.58	3	Viverricula_indica	117.75	-8.58	2
Hippopotamus_amphibius	27.08	-25.25	2	Ondatra_zibethicus	-169.58	56.58	3	Viverricula_indica	43.42	-11.75	2
Hippopotamus_amphibius	25.58	-33.42	4	Ondatra_zibethicus	-132.75	55.58	3	Vombatus_ursinus	148.08	-42.58	5
Hippotragus_equinus	31.83	-2.33	3	Ondatra_zibethicus	-164.25	54.75	3	Vulpes_vulpes	146.42	-38.92	3
Hippotragus_equinus	30.83	-17.92	4	Ondatra_zibethicus	-125.58	49.75	3	Vulpes_vulpes	37.75	55.58	2
Hippotragus_equinus	15.75	-18.92	3	Ondatra_zibethicus	-63.00	49.42	3	Vulpes_vulpes	-4.58	54.25	2
Hylobates_lar	100.92	12.75	2	Oreamnos_americanus	-149.75	60.42	3	Vulpes_vulpes	40.25	51.08	2
Hystrix_brachyura	-3.92	50.75	4	Oreamnos_americanus	-124.08	48.75	0	Vulpes_vulpes	-125.58	49.75	3
Hystrix_cristata	-2.08	53.08	5	Oreamnos_americanus	-135.75	57.92	3	Vulpes_vulpes	-1.42	50.75	2
Hystrix_cristata	14.33	37.58	5	Oreamnos_americanus	-153.08	57.58	3	Vulpes_vulpes	-63.00	49.42	3
Isoodon_auratus	125.58	-24.75	2	Oreamnos_americanus	-135.08	57.08	5	Vulpes_vulpes	-123.00	48.58	3
Isoodon_obesulus	167.83	-47.08	0	Oreamnos_americanus	-123.42	47.75	3	Vulpes_vulpes	-70.58	41.42	3
Kobus_kob	-16.58	12.42	2	Oreamnos_americanus	-110.58	47.42	3	Vulpes_vulpes	143.83	-37.58	3
Kobus_leche	26.25	-32.75	3	Oreamnos_americanus	-109.33	46.75	3	Zaedyus_pichiy	-72.08	-36.75	3
Kobus_leche	30.83	-17.92	3	Oreamnos_americanus	-110.25	46.08	3	Zaedyus_pichiy	-72.08	-36.75	3
Lama_guanicoe	7.00	48.25	4	Oreamnos_americanus	-112.42	45.75	3	Zaedyus_pichiy	-72.08	-36.75	3
Lama_guanicoe	0.58	42.58	4	Oreamnos_americanus	-109.58	45.08	3				

PART II

Predicting biological invasions in a changing world

CHAPTER 2.1

Will climate change increase the risk of plant invasions into mountains?

Blaise Petitpierre[1], Keith McDougall[2], Tim Seipel[3], Olivier Broennimann[1], Antoine Guisan[1] & Christoph Kueffer[4]

[1] Department of Ecology & Evolution, University of Lausanne, 1015 Lausanne, Switzerland

[2] Department of Environment and Climate Change, Queanbeyan, NSW, Australia

[3] College of Agriculture, 728 Leon Johnson Hall, Montana State University, Bozeman, USA

[4] Institute of Integrative Biology, ETH Zürich, Universitätstrasse 16, 8092 Zürich, Switzerland

This paper is in preparation for *Global Change Biology*

My contribution to the project: I designed the statistical approach, ran the niche modeling analyses and wrote the initial draft of the paper

ABSTRACT

Mountain ecosystems are not yet badly affected by invasions of non-native plants. However, many invasive plants of lowland ecosystems currently show a sharp elevational distribution limit at 1000 to 1500 m asl, which is likely linked to climate. In a warming climate these species may move upwards and in the future also threaten mountain ecosystems.

We modeled the realized niche of 49 plant invaders through species distribution modeling (SDM) to predict their current and potential future distribution in two contrasted mountain areas, the Swiss Alps (CH) and the Australian Alps in News South Wales (NSW). Both areas are characterized by different topo-climatic conditions.

Optimal suitability for plant invaders is predicted to shift from lowland to montane and subalpine elevations in CH, whereas the shift is less pronounced in NSW, montane and subalpine elevations being currently suitable, with a slight shift predicted towards the alpine level. We examined this contrasted distributional response in regards to species traits and discuss the implication for future biodiversity in mountain areas.

INTRODUCTION

Non-native plant species richness in mountainous areas is generally low, especially at high elevations (Seipel *et al.*, 2012, McDougall *et al.*, 2011*a*) where only few species become non-native and negatively affect ecosystems and biodiversity (Kueffer *et al.*, 2014, McDougall *et al.*, 2011*b*). Several possible explanations have been proposed to explain these patterns, which contrast strongly with non-mountain ecoregions (Pauchard *et al.*, 2009). It could be that elevational richness patterns of non-native species might be shaped by the same factors as those of native species: energy, productivity, and/or water availability are reduced with elevation (Alexander *et al.*, 2011, Jakobs *et al.*, 2010, Rahbek, 1995, Romdal and Grytnes, 2007, McCain, 2007), leading to low richness of both native and non-native species. Alternatively, it has been proposed that introductions of non-native species have predominantly occurred at low elevation first, resulting in a lag for invasion at high elevation (Becker *et al.*, 2005, Alexander *et al.*, 2011). Less intensive land use and human activity at high elevations may reduce invasion opportunities because of lower propagule pressure and/or higher resistance of undisturbed habitats to invasion (Jakobs *et al.*, 2010, Marini *et al.*, 2009, Pauchard and Alaback, 2004, Kueffer *et al.*, 2014, Parks *et al.*, 2005). Adaptive evolution may also play a role in this pattern, as recently proposed by Hufbauer *et al.* (2012). Populations adapted to human-altered habitats in the lowland areas of the native range are more likely to have been associated with human transport mechanisms, enhancing the likelihood of transport to a novel range where it can succeed in similar lowland, human-altered habitats. Such a preponderance of introduction of lowland species would lead to the absence of non-native species at higher elevation except those with a broad climatic niche (‘directional ecological filtering’, Alexander *et al.*,

2011). Indeed, species pre-adapted to a mountain climate are conspicuously lacking from high elevation nonnative floras worldwide (McDougall *et al.*, 2011*a*). Consequently, climate seems be an important factor limiting the spread of lowland non-native floras into high elevation mountain ecosystems (Jakobs *et al.*, 2010, Seipel et al., in revision, Kueffer *et al.*, 2014, Trtikova *et al.*, 2010, Marini *et al.*, 2009).

However, this situation could change. First, with time, propagule pressure and habitat disturbances might increase in mountainous regions due to economic development, in particular the development of tourism (McDougall *et al.*, 2005, McDougall *et al.*, 2011*b*), leading to the possibility of transport of species directly between mountainous regions. Second, it is also feared that climate change might reduce the climatic limitation of current non-native species distributions and facilitate invasions into mountains (Barni *et al.*, 2012, Pauchard *et al.*, 2009, Kueffer, 2010, Walther *et al.*, 2009). Indeed, there is some evidences that human, animal, and plant diseases will move uphill with climate change and threaten mountain livelihoods, ecosystems, and biodiversity (Walther *et al.*, 2009, Kurz *et al.*, 2008, Benning *et al.*, 2002, Hay *et al.*, 2002). To our knowledge whether plant invasion risks will increase in mountain areas has never been systematically and quantitatively investigated although impacts on ecosystems and economy might be as important as those from climate change.

The response of a specific non-native species to climate change is expected to vary between areas of introduction for three reasons. First, different climate factors might interact in shaping species distribution limits. For instance, aridity, more than temperature, seems to act as a limiting factor towards higher elevation in dry areas (e.g. in Hawaii, Jakobs *et al.*, 2010, Juvik *et al.*, 2011). Depending on the limiting factors, species distribution might respond differently to climate change because not all climate factors will change with the same magnitude and direction (Crimmins *et al.*, 2011, McCain and Colwell, 2011). Second, impacts of climate change will not be homogenous all around the world and some mountain chains will be more affected than others (Engler *et al.*, 2011). And third, elevation gradient can be shorter or smoother from one area to another offering different distances from the elevation “dead-end” (i.e. the top of the gradient)

Anticipating biological invasions was shown to be more efficient than post introductions control and eradication efforts (Leung *et al.*, 2002). Accordingly, good anticipation will allow a better conservation management of the biodiversity hotspots and the pseudo-islands that are mountain chains. To do this, species distribution modeling (SDM) based on environmental niche modeling allows identifying potential suitable habitats for non-native species in geographic space (Guisan and Zimmermann, 2000). SDM relates species distribution to environmental variables (generally available as GIS layers) and fits the Hutchinsonian realized species niche (*sensu* Soberon, 2007; i. e. also integrating dispersal limitations). To depict a reliable potential distribution, this approach relies on two major assumptions. First biotic interactions are assumed to be the same in the calibrated and projected areas. Second, species are assumed to have reached their niche equilibrium, i.e. that hey have colonized all possible suitable environmental conditions.

Although some realized niche shifts have been detected between native and invaded ranges for some plant species (e.g. Broennimann *et al.*, 2007, Gallagher *et al.*, 2010), most Holarctic invaders show limited niche expansion (Petitpierre *et al.*, 2012) supporting for the possibility to predict their invaded range from the native range. Moreover, pooling all available distribution data allows building more accurate SDM, even – at least for the time considered - for niche shifting species (Broennimann and Guisan, 2008). Additionally, provided that climatic scenarios are available, SDMs allow assessing the response of plant distributions to climate change (e.g. Engler *et al.*, 2011).

Scale and resolutions also matter for building SDMs and deriving projections, especially in mountainous landscape (Randin *et al.*, 2009*a*, Seipel *et al.*, 2012). On the one hand, SDM at a global scale will identify species niches more comprehensively, if the full range of suitable environmental conditions for the species is not represented within the focal study extent (i.e. truncated niches; Thuiller *et al.*, 2004). However, finer resolution environmental and distributional data over large-enough geographic extents to encompass full climatic niches of species are rarely available, requiring the use of coarser resolution data for these more global SDM approaches. With such coarse grain SDMs, outlying values important to the species may not be reflected by an average value within one coarse spatial cell and predictions may differ substantially from SDM calibrated at a finer, local scale (Randin *et al.*, 2009*a*). Thus, using finer SDM over restricted extents may allow catching more proximal values of environmental factors controling distributions than coarser SDM over larger extents but they may fail to catch the entire species' realized niche. This is why modeling in a combined way at both local and global scales would allow keeping wide biogeographical perspective on the niche and distributions, while adding the finer environmental perspectives on local environmental requirements and fine distributions of species. Such development looks particularly useful for predicting potential distributions of non-native species in mountainous landscapes.

In addition to SDM, understanding which traits and characteristics make some non-native species more invasive than others is another possibility to predict and understand invasions success (Van Kleunen *et al.*, 2010 but see Thompson and Davis, 2011). Combining species traits to SDM analyses should allow a wider understanding of mechanisms explaining the tremendous colonization success of some invaders, with potential to further (Moles *et al.*, 2008, Gallagher *et al.*, 2010, Hulme and Barrett, 2013).

In this paper, we take the challenge of assessing whether mountain areas will be at increased risk of plant invasions under climate change, using a combined approach linking coarse data models at global scale with fine distribution data and environmental maps over local mountain ranges. We modeled potential distributions for 49 invasive plant species using SDM in two climatically contrasted mountainous areas, at both global and local scales. We chose one study system - Switzerland - where temperature is assumed to be the most stressful factor, in particular at the higher edge of the elevation gradient and one study system - New South Wales (in Australia) - with a gentler elevation gradient, where drier climate at the lower edge may have an important role in shaping species

distributions. We assessed the magnitude of the elevation shift for invasive plant species in mountains using the predictions of correlative SDM for seven different climatic scenarios. We examined the links between species traits and their distribution along an elevation gradient under current and future climates.

The specific questions we aimed to answer in this paper were:

1) Will mountain become more sensitive to plant invasions under global warming?

2) Will global warming affect two different mountain ranges with the same magnitude?

3) What kind of species will be able to invade mountains?

METHODS

The study areas

Switzerland (CH, total area 41,285 km^2) is characterized by a mountainous landscape with a large elevation gradient (between 192 and 4634 m asl, Fig. 1a). The Alps are the most important mountain chain occupying 60% of the country area at the center, eastern and south-eastern sides. On the western side, the Jura is smaller and less elevated (10% of the country area and rising up to 1680 m asl). Western European broadleafs forests and European-Mediterranean montane forests are the two ecoregions found in CH (Olson *et al.*, 2001). CH was classified into 4 elevation classes: lowland (192 - 800 m), montane (800 - 1500 m), subalpine (1500 - 2200 m) and alpine (2200 - 3100). Higher altitudes were not included in the analyses because climate becomes unsustainable for most plant species.

New South Wales (NSW, total area 809,444 km^2) is a State in the east of Australia, containing part of the Australian Alps in the far south-east. It contains the highest peak of Australia (2228 m asl) but is also characterized by an important gradient in precipitation from west to east and 6 ecoregions: Eastern Southeast Australia temperate forests, Australian Alps montane grasslands, Eastern Australian temperate forests, Southeast Australia temperate savanna, Murray-Darling woodlands and mallee and Eastern Australia mulga shrublands (Olson *et al.*, 2001). NSW was classified into 5 elevation classes (Fig. 1b). Due to a strong aridity gradient present in this area, we divided the lowland area wet and dry: dry lowland (0 - 700 m and current annual precipitation < 500 mm), wet lowland (0 - 700 m and current annual precipitation > 500 mm), montane (701 - 1400 m), subalpine (1401 - 1800 m) and alpine (1801 - 2228m). As no species were present in the Simpon's desert ecoregion, this area was not included in the analyses. Climatic differences between CH and NSW for temperature and precipitation along elevation gradient are illustrated by Fig. S1

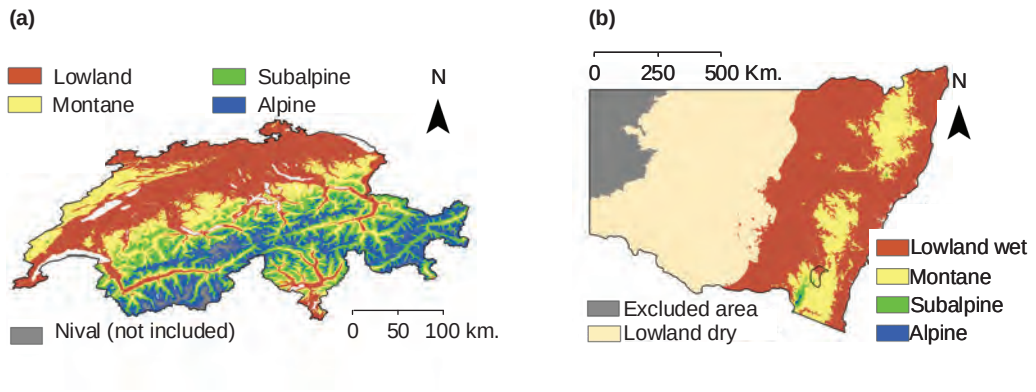


Figure 1: Study area and distribution of elevation level in Switzerland (CH, a) and New South Wales (NSW, b). The elevation gradient was divided into lowland (192 - 800 m in CH, 0 - 700 m in NSW), montane (801 - 1500 m in CH, 701 - 1400 in NSW), subalpine (1501 - 2200 m in CH, 1401 - 1800 m in NSW) and alpine levels (2201 - 3100 m in CH, 1800 - 2228 m in NSW). Note that areas with extreme climatic conditions were removed from the study and that lowland in NSW is split into wet (> 500 mm of annual precipitation) and dry areas (< 500mm of annual precipitation)

Species data

The national black and watch lists were used as a reference to choose the invasive species in CH (infoflora: www.infoflora.ch). Only terrestrial plant species for which enough precise georeferenced occurrence data at both local and worldwide scale were used (Tables S1 and S2). We screened the following local databases: infoflora (www.infoflora.ch), relevés from the Mountain Invasion Research Network (MIREN available in Seipel, 2011) and from previous work related to invasive plants along elevation gradients (Becker *et al.*, 2005).

Data for NSW invaders were obtained from relevés stored in the Atlas of NSW Wildlife (http://www.bionet.nsw.gov.au/; accessed 5 October, 2013). Twenty-four species were selected that occurred in or close to the Australian Alps, had not established throughout the Alps, and had sufficient records for robust SDM. Data for five additional species that had been included for CH were also obtained ((*Ailanthus altissima*, *Ambrosia artemisiifolia*, *Lonicera japonica*, *Prunus laurocerasus* and *Robinia pseudoacacia*, Tables S3 and S4).

In essence, the species chosen for SDM in both CH and NSW are a random selection from the pool of non-native species present in these areas. There is no prior reason to expect that any will be especially suited to high altitude environments nor that they will respond to climate change in particular directions.

Both local databases were completed with data from the Global Biodiversity Information Facility (*data.gbif.org*, accessed between August and September 2012) for the SDM at global scale. We retained occurrences with a known accuracy higher than 2.5 km and 250 m for global and local SDM respectively.

Multi-scale SDM framework

We fitted SDM at two scales: global and local. First, global SDM with worldwide species distributions associated with coarse climatic variables were built. This allowed capturing the widest possible climatic niche of species. This global SDM based on the climatic niche was used in two ways: (i) to predict the current and future worldwide species' distribution build under scenarios of climate; and (ii) to further select the pseudo-absences to be used in the local SDM (Gallien *et al.*, 2012). The local SDM were calibrated at a much finer resolution and included climate but also disturbance variables affecting species distribution at short distances.

The global SDM

Eight climatic variables were used at the global scale (Table 1). These variables are known to be robust to SDM transferability (Petitpierre et al. *in prep*) and were available on the Worldclim database at 30 Arc seconds (about 1km at the equator, Hijmans *et al.*, 2005) ,except aridity and aridity seasonality. Aridity data were available on the CGIAR-CSI website (http://www.cgiar-csi.org) and seasonality was assessed through the variance of the 12 monthly aridity means.

It is important to define an ecologically-relevant study area where SDM are calibrated (Barve *et al.*, 2011). Inadequate extents may result in truncated SDM or over-optimistic evaluation (Guisan and Thuiller, 2005, Chefaoui and Lobo, 2008, Barve *et al.*, 2011, Acevedo *et al.*, 2012). For this purpose, we defined the study area for each species as the ecoregions covered by at least one occurrence. Ecoregions are geographical units defined by homogeneity between climates, geology, fauna and flora (Olson *et al.*, 2001).

The local SDM

For each species, we used georeferenced data of the local databases with accuracy equal to or higher than 250 m. Within each study area, we defined the extent in the same way as for the global SDM (i.e. ecoregions with at least one species occurrence). Within this extent, pseudo-absences were selected inversely proportionally to the suitability provided by the global SDM (as in Gallien *et al.*, 2012, but with a linear weight). Three climatic variables and two disturbance variables were used at a 100 m and 250 m resolutions for CH and NSW respectively (Table 1). For CH, climate variables were used from Zimmermann and Kienast (1999) whereas for NSW they were provided by the NSW Office of Environment and Heritage (Christopher Simpson for NSW, personal communication).

Table 1: Variables description

Abbreviation	Scale	Variable description
G1	Global	Annual Mean Temperature
G2	Global	Temperature Seasonality
G3	Global	Mean Temperature of Warmest Quarter
G4	Global	Mean Temperature of Coldest Quarter
G5	Global	Precipitation Seasonality
G6	Global	Precipitation of Warmest Quarter
G7	Global	Aridity
G8	Global	Aridity seasonality
L1	Local	Annual Mean Temperature
L2	Local	Temperature seasonality
L3	Local	Annual precipitations
L4	Local	Density of urban area
L5	Local	Distance to the closes river,lake or shore

Modeling techniques

For both global and local SDM, we randomly sampled 10 000 pseudo-absences in the calibration extent of each species. To reduce spatial autocorrelation due to sampling bias, occurrences were disaggregated to keep a 10 km minimal distance between two occurrences in the global SDM and 1 km in the local SDM. Data were split into calibration (70%) and evaluation (30%) datasets. We applied an ensemble modeling approach by averaging three modeling techniques: GLM with polynomial quadratic coefficients preceded by a stepwise selection based on the Bayesian Information Criteria (BIC) (McCullagh and Nelder, 1983), Boosted regression tree (with 5000 trees and 7 degrees of interaction depth, Friedman *et al.*, 2000) and MAXENT (with default parameters, Phillips *et al.*, 2006). GLM and GBM were done with the BIOMOD package (Thuiller *et al.*, 2009) and MAXENT with dismo package (Hijmans *et al.*, 2005) on the R software (R, 2005). Variable contribution was assessed through randomization (see the documentation of BIOMOD). Suitability and variable contribution were averaged between each technique. The resulting averaged suitability predictions were evaluated with the Area Under the Curve of a Receiver Operating Characteristics (Roc AUC, Zweig and Campbell, 1993), True Skill Statistics (TSS, also known as corrected Kappa, Allouche *et al.*, 2006) and two “presence-only” evaluators, the Boyce index (B, Hirzel *et al.*, 2006) and the rate of correctly predicted presences with binarized prediction, also called sensitivity. Presence-only evaluators are particularly relevant in the case of invasive species because they are ongoing processes, and observed distributions may not reflect their distributional equilibrium. Apparent absences may thus not be reliable for model evaluation. For clarity, we only present AUC and Boyce in the manuscript (but see Tables S1 - S4 for the other evaluators). Continuous probabilities were binarized using the threshold providing the best TSS. For each species, we assessed the proportion of the study area covered by the potential distribution in relation to elevation categories and measured the highest potential elevation. The whole procedure was replicated 10 times per species and averaged.

Additionally, climate analogy between the calibration and the projection areas was assessed by calculating the Multivariate Environmental Similarity Surfaces (MESS, Elith *et*

al., 2010) at each site of the study area and for each climatic scenario. Extrapolating SDM predictions to non-analog climate may provide spurious and thus unreliable results (Fitzpatrick and Hargrove, 2009, Mandle *et al.*, 2010, Peterson, 2011). We performed the analyses with and without considering climate analogy but since it did not effect our main conclusions about higher elevations, we present only scenarios where non-analog sites are considered as unsuitable (but see Fig. S2 for results without considering climate non-analogy).

Spatial projections under future climate

Scenarios for global SDM were directly available on the website from the Research Program on Climate Change, Agriculture and Food security (http://www.ccafs-climate.org/, Ramirez and Jarvis, 2008). Climatic scenarios for variables G7 and G8 (Table 1) were derived following Zomer *et al.* (2008) using bioclimatic variables of Worldclim. We applied 6 different future climatic scenarios based on two emission scenarios of the IPCC 4th assessment report (A1B and A2, IPCC, 2007), two reference dates (2030 and 2070) and three GCMs (HADCM3, ECHAM5 and CCSM3, Tables S5 to S8). For local SDM in CH, RCMs were available at a 1 km resolution (Niklaus Zimmermann, personal communication). Climate anomaly was bilinearily interpolated at a 100 m resolution. For NSW, RCM were available at a 250 m resolution (Christopher Simpson, personal communication).

Principal component analysis (PCA)

To visualize how climate change will affect species climatic space, we built a PCA calibrated on the pool of both native and invasive areas using climatic data of the global SDM (referred as PCA_{env} in Broennimann *et al.*, 2012), with the library ade4 in R, (Dray and Dufour, 2007). Note that in CH along each variable, percentile above the 95th were removed as these conditions are located in extreme alpine and nival climates that are not relevant for plant species distributions. Species' global distributions were then projected along these two components and the frequencies of occurrence were smoothed in a gridded environmental space (resolution R = 200) using a kernel density function in the same way as Broennimann *et al.* (2012). This approach allows visualizing, delimiting and piling up species niches and study extents in the PCA climatic space (Petitpierre *et al.*, 2012).

Species traits

For each species, we collected a serie of traits and characteristics (Table S9) from two databases: Flora indicativa (Landolt, 2010) and Biolflor (Kuhn *et al.*, 2004). Although some elements of these tables are not strictly species traits (e.g. Landolt's environmental indicators), we referred to all of them as traits for clarity. These traits were used to

determine if the species invading the two different study areas had different characteristics and if traits can explain predicted elevation shifts towards higher altitude during global warming.

We began with testing whether CH species had different traits to NSW species with two sided random permutation test when variables were (semi-)quantitative and non-parametric Chi-Square tests when variables were qualitative.

We further measured whether species traits and characteristics could be associated with their maximal altitude. We tested three types of responses: 1) current observed maximal altitude, 2) current predicted maximal altitude and 3) shift between predicted maximum altitude under current conditions and predicted maximum altitude under climate conditions in A1 2070 scenarios. Future maximal predicted altitudes were averaged for the two A1 2070 scenarios (scenarios 6 and 7 on Tables S1 to S4). In the model of the predicted altitude shift, we added the elevation potential (i.e. the difference between the current predicted highest altitude for each species and the maximum altitude of the gradient in the region) as an explanatory factor. For both NSW and CH species pools, we tested these associations for global and local SDM outputs. As there were many traits compared with the number of species, we performed a hierarchical stepwise approach to select important traits for the responses and avoid model over-parametrization. First, for each trait we built univariate GLM including only one trait and compared the BIC of the obtained model to a model with only the intercept (null model). If the model including the trait did not have a better BIC than the null model, the trait was not considered further. After this first selection, all combinations of the retained variables were examined and the final combination providing the best BIC was used as the final model. Multimodel inference and variable selection was done in R using the library MuMIn (Barton, 2012)

RESULTS

Suitability in CH by global SDM

SDM show overall good accuracy with an average AUC of 0.88 ±0.05 and a B of 0.8 ±0.22 (Table S1). Species show neither poor AUC (below 0.7) nor poor B (below 0). The most important variables for invasive species in CH are temperature variability, temperature of the coldest quarter and aridity index (Fig. 2). Under current conditions, the average suitable area represents 84 ±22%, 57 ±35%, 17 ±23% and 3 ±6% of the total area in lowland, montane, subalpine and alpine zones respectively (Fig. 3 and 4). For most species potential highest elevations (1843 ±473 m) are significantly higher than the ones observed (=1317 ±497 m, t-test $P < 0.001$, Fig. 5). The difference between potential and actual observed highest elevation varies between -74 m for *Solidago canadensis* and 1572 m for *Phytolacca americana* with an average of 546 ±413 m (Table S1). Climatic scenarios for the future show a decrease of potential suitability at the lowland level, no changes at the montane level and an increase at subalpine and alpine levels (Fig. 3 and 4). The

average highest elevation increases with climate warming (Kruskall-test, $P < 0.001$), reaching 2430 ±474 m for scenario Hadcm1 A1B 2070 (Fig. 5). For example, when results from both A1b scenarios are averaged, the difference between current and future potential elevation varies between -79 m for *Lonicera japonica* and 995 m for *Prunus laurocerasus* with an average of 579 ±286 m (Table S1). In the future, up to 7 species will experience non-analog climates at the lowland level (Fig. S3). Spatial projections are available for each species and each climatic scenario in Fig. S4.

Suitability in CH by local models

SDM show good overall accuracy, with an average AUC of 0.91 ±0.06 and a B of 0.72 ±0.24 (Table S2). Species show neither poor AUC nor poor B (below 0). Average temperature is the most important variable driving species distributions in CH (Fig. 2). Under current conditions, only the lowland level shows a large proportion of suitable area; higher vegetation levels remain largely unsuitable for the majority of the species with suitable areas representing 37 ±19%, 4 ±6%, 1 ±4% and 0 ±0% of the total area at lowland, montane, subalpine and alpine levels respectively (Fig. 3 and 4). On average, lowland and montane levels are predicted by climatic scenarios to experience the most dramatic increases in suitability in the future. However, according to the warmest scenarios, suitability is predicted to decrease at the lowland level because no longer analog (but see Fig. S2 when SDMs are extrapolated to novel climates). For most species potential highest elevations (1331 ±473 m) are not significantly different from the ones currently observed (1317 ±497 m, t-test $P = 0.91$, Fig. 4), varying between -562 m for *Reynoutria japonica* and 602 m for *Cornus sericea* with an average of 16 ±341 m (Table S2). Climatic scenarios have a significant effect on the maximal predicted elevation of distributions (Kruskall-test, $P < 0.001$), with potential highest elevations reaching 2170 ±340 m for scenario Hadcm3 A1b 2070 (Fig. 4). When results from both A1b 2070 scenarios are averaged, the difference between current and future potential elevation varies between 252 m for *Sedum spurium* and 979 m for *Bunias orientalis* with an average of 722 ±179 m (Table S2). The average spearman correlation between local and global SDM predictions of the highest elevations under the 7 scenarios is 0.57 ±0.13. Spatial projections are available for each species and each climatic scenario in Fig. S5.

Suitability in NSW by global models

On average, models have similar evaluations to CH species, with good AUC (0.89 ±0.04) and good Boyce index (0.95 ±0.05). Species show neither poor AUC nor poor B. The most important variables are the same as for CH, i.e. temperature variability, temperature of the coldest month and average annual temperature (Fig. 2). However, temperature seasonality is more important for species in NSW than in CH (t-test $P < 0.001$, Fig. 2). Under current conditions, suitability increases with elevation, reaching a peak at the

montane and subalpine levels. Potential suitable area represents 6 ±11%, 33 ±26%, 61 ±30%, 63 ±38% and 40 ± 42% of the total area at dry lowland, wet lowland, montane, subalpine and alpine levels respectively (Fig. 3 and 4). Climatic scenarios predict a decrease of suitability at lowland and montane levels and an increase of suitability at subalpine and alpine levels (Fig. 3 and 4). For most species, current predicted highest elevations (1789 ± 271 m) are significantly higher than the highest elevation currently observed (= 1287 ± 292 m, t-test P < 0.001, Fig. 5), varying from -46 m for *Pinus radiata* to 1039 m for *Ilex aquifolium* On the average, climatic scenarios have a significant effect on the maximal predicted elevation of distributions (Kruskall-test, P < 0.001), reaching 1969 ± 190 m (Fig. 5) with the warmest scenario echam5 A2 2070. For example, when results from both A1b 2070 scenarios are averaged, the difference between future and current predicted elevation varies between -70 m for *Chondrilla juncea* and 432 m for *Ambrosia artemisiifolia* with an average of 157 ± 144 m (Table S3). In the future, climate becomes non-analog for up to 12 species, essentially at lowland and on the warm and dry western part of NSW (Fig. S3). Spatial projections are available for each species and climatic scenario in Fig. S4.

Suitability in NSW by local models

On average, models have excellent AUC (0.95 ±0.05) and good B (0.76 ± 0.04). The most important variables in driving species distributions in NSW are temperature, precipitation and temperature variability (Fig. 2). All variables, except distance to rivers and lakes, are significantly different from the CH local SDM (Fig. 2). Under current conditions, suitability reaches a peak at the montane level with average proportion of suitable area of 2 ±5%, 18 ±15%, 43 ±32%, 35 ±39% and 28 ±40% (Fig. 3 and 4). Under current conditions, arid lowland remains unsuitable for almost all species whereas higher levels are suitable for a most of the species (Fig. 3 and 4). Climatic scenarios predict a decrease of suitability at lowland and montane levels, a slight increase of suitability at subalpine level and stable suitability at alpine level (Fig. 3 and 4). Current predicted highest elevations (1628 ±447 m) are significantly different from the ones observed. (t-test P = 0.001, Fig. 5), varying from -80 m for *Convolvulus arvensis* to 1030 m for *Prunus laurocerasus*. On the average, climatic scenarios have no significant effect on the maximal elevation of distributions (Kruskall-test, P = 0.98, Fig. 5). For example, when results from both A1b scenarios are averaged, the difference between current and future average predicted highest elevation varies between -979 m for *I. aquifolium* and 523 m for *A. artemisiifolia* with an average of 50 ±262 m (Table S5). The average spearman correlation between local and global SDM predictions of the highest elevations under the 7 scenarios is 0.55 ±0.03. Spatial projections are available for each individual species and each climatic scenario in Fig. S5.

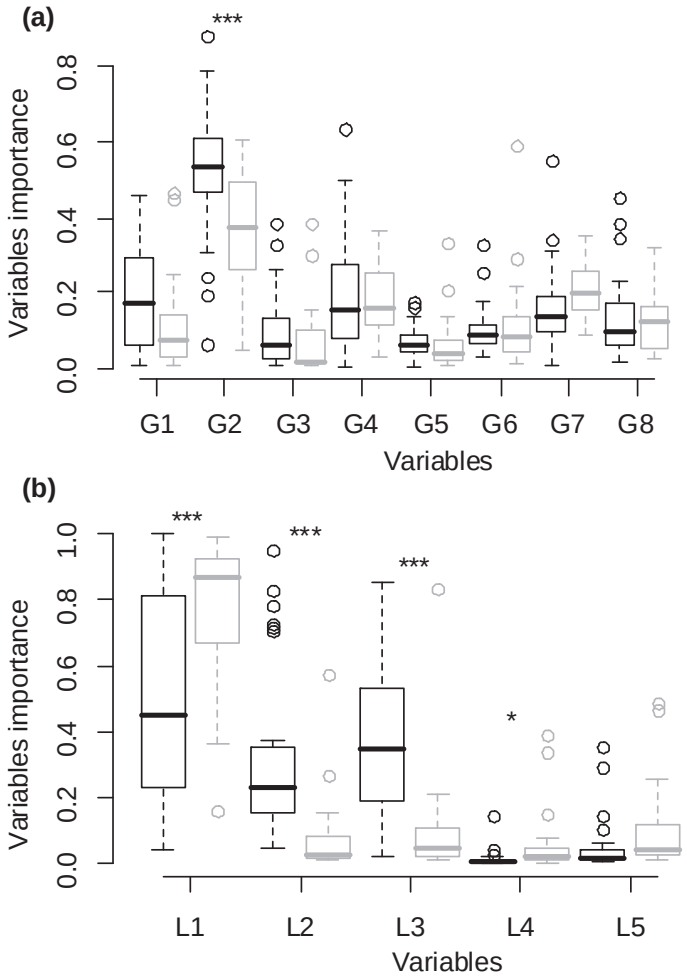


Figure 2: Variables importance (abbreviated as in Table 1) for species invasive in New South Wales (black) and Switzerland (grey) in global (a) and local (b) models. Stars represent the significance of the difference between CH and NSW value for each variable assessed with a t-test

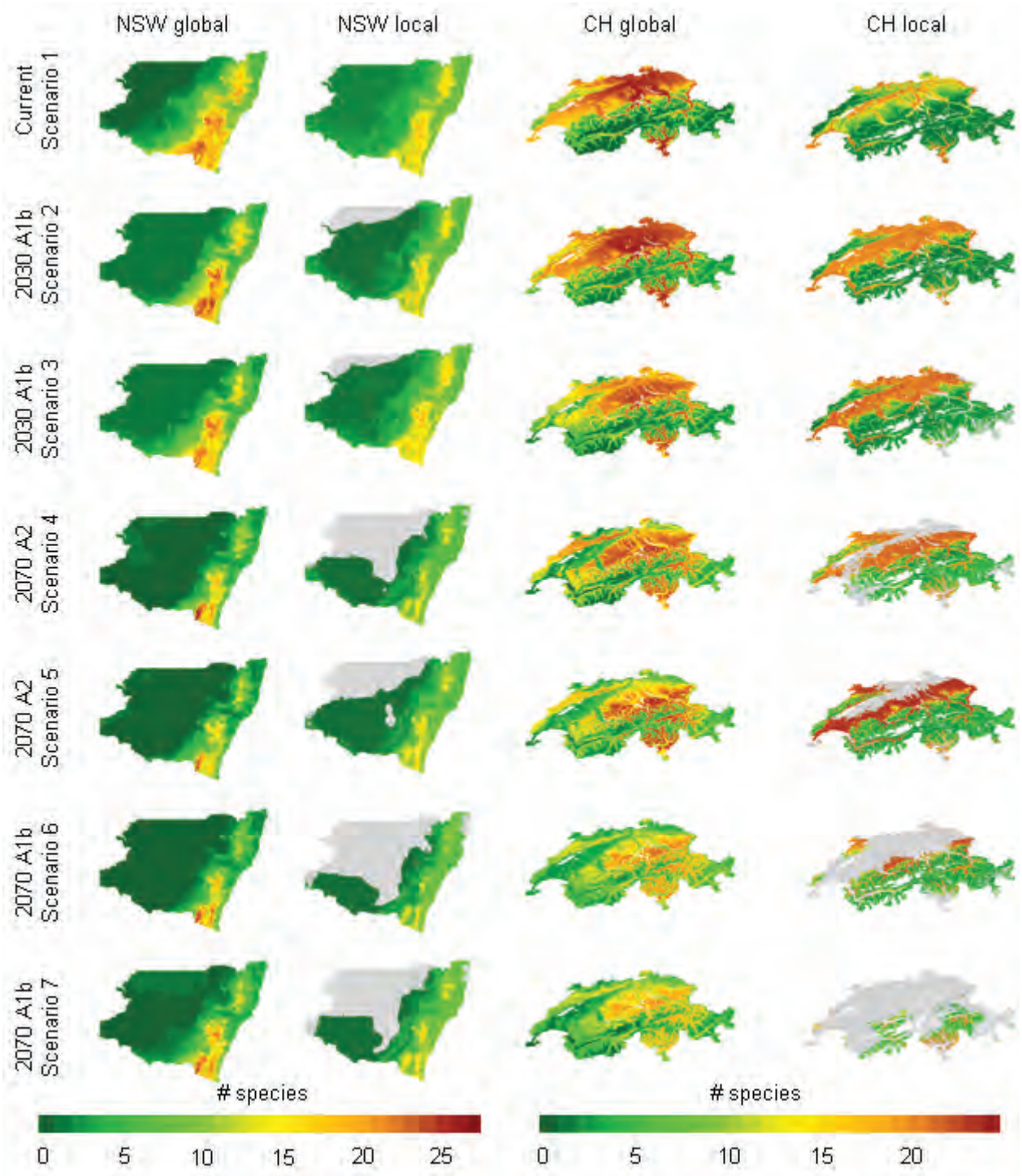


Figure 3: Piling of binarized potential species distribution provided by local and global models in Switzerland and New South Wales under seven different climatic conditions (see table S6 to S9 for a detailed description of the scenarios). Colors represent the potential number of species (# species). Non-analog climate is depicted in grey.

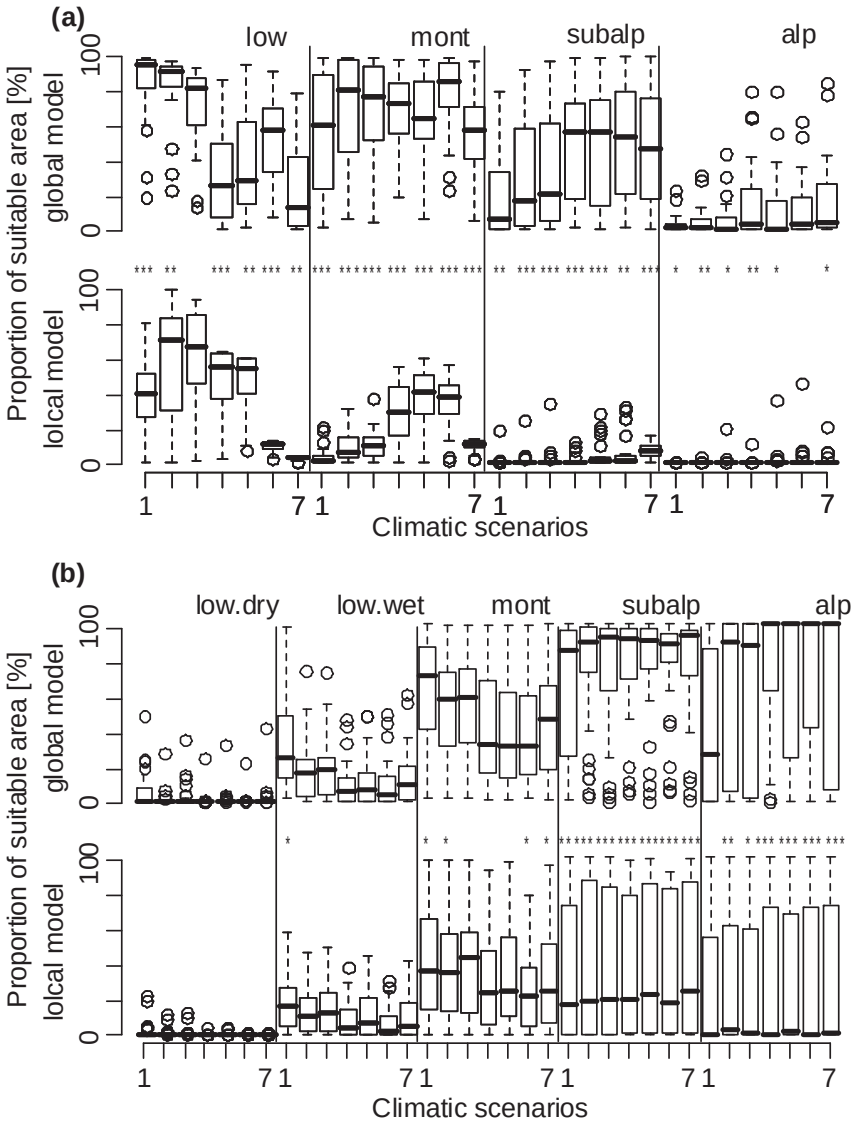


Figure 4: Proportion of suitable area predicted by global and local models for the seven climatic scenarios (ranked from 1 to 7, following table) at each vegetation level (lowland, montane, subalpine and alpine) in Switzerland (a) and New South Wales (b)

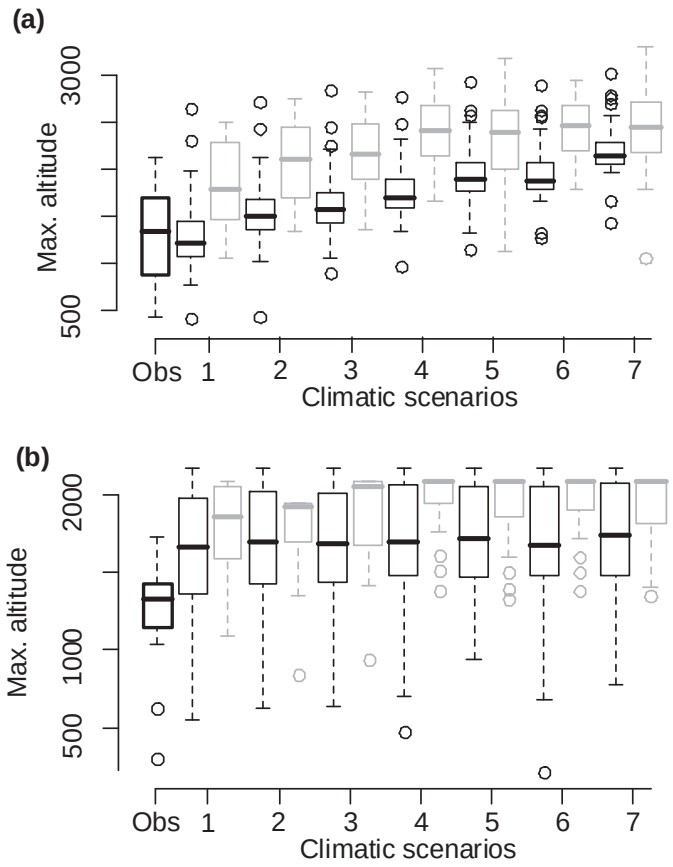


Figure 5: Highest observed (Obs) and predicted altitude following different climatic scenarios (ranked from 1 to 7) predicted by local (black) and global (grey) models in Switzerland (a) and New South Wales (b).

PCA

The two first axes of the PCA summarize 91% of the climate variation in CH and 83% in NSW (Fig. 6). The PCA scores show that invasive species occupy in CH the warmer portion of conditions present in the study area and seem thus mostly limited by cold temperatures. On the contrary in NSW, although the pattern is less contrasted than in CH, only the aridity gradient seem to a limiting factor. The PCA also shows that climatic scenarios have different influences in CH from NSW. In CH, in addition to a shift toward warmer and dryer climates, a contraction is predicted along precipitation gradients (i.e. the variability of climatic conditions is reduced). On the other hand in NSW, climate is predicted to shift towards dryer and hotter conditions. Species distribution will be limited by high temperature in NSW, in addition to aridity.

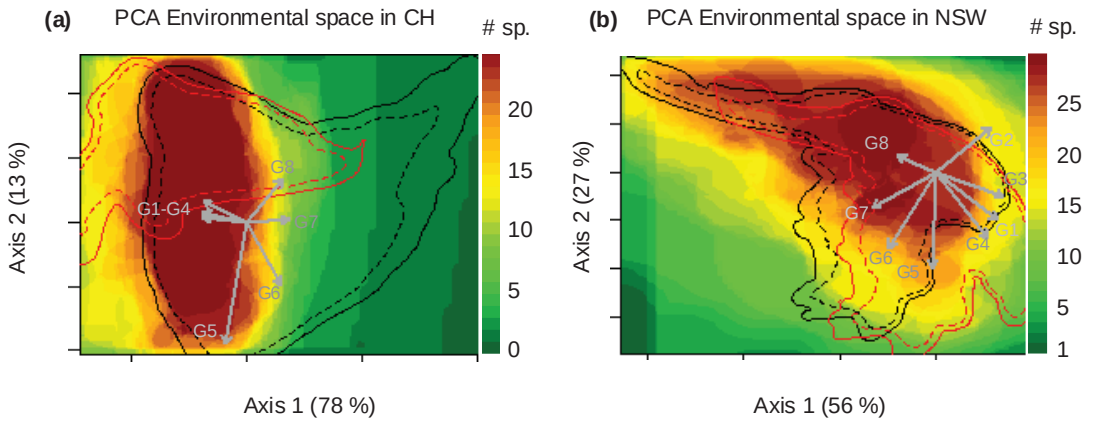


Figure 6: Environmental space depicted by the two first axes of a principal component analysis (PCA) calibrated on Switzerland (CH) and New South Wales (AU). Species niches were projected and piled into this environmental space such as each environmental condition is characterized by the number of species for which it is suitable. Correlation between variables and components are represented by grey arrows and each number correspond to variables in Table 1. Black lines represent conditions in the study area under current conditions and red lines conditions under future hadcm3 A1 2070 scenarios. Dashed lines represent the distribution till the 75th percentile of all climatic conditions in the study area.

Species traits

Significant differences between species pools in CH and NSW were found in Landolt's temperature (4.27 vs 3.85; t-test P = 0.040) and soil moisture variability (1.92 vs 1.42; t-test P = 0.027), life strategies (CH species are more competitive, Chi-square test P = 0.019), dominance (4.04 vs 1.62, t-test P < 0.001), number of occurrences (t-test P < 0.001) and the number of species present in southern temperate areas (15 vs 20 species, Chi-square test P = 0.044, Table S9).

Species traits explained between 25% and 86% of the variation in species' maximal altitudes (observed or predicted, Table 2, Fig. S6). In NSW, species maximal altitude is associated with various traits, but not with temperature, while soil moisture predicts the observed maximum elevation and the shift of predicted maximum altitude (Table 2). In global SDM, stress-tolerant competitors are predicted to have lower SDM predicted maximal altitudes whereas competitive ruderals and species with a wider elevation potential will expand their altitudinal range (Table S10). In contrast, the shift in altitudinal response predicted by local SDM is associated with the potential of elevation of species and their ability to keep their leaves in winter. Quite interestingly, the traits responsible for maximal elevation show a completely different pattern in CH. Species' resistance to cold is associated with every measured response, except with the shift in current vs. future predicted altitude which was only correlated with elevation potential. Additionally, species with rosette, diploidy and growing nutrient rich soils seem to predict well the maximal altitudes in this range (Table S10).

Table 2: GLMs assessing the relationship between three types of maximal elevations and species' traits. Actual maximal elevations, maximal predicted elevation by SDMs. The significant species' traits remaining after the stepwise selection and the explained deviance of the models are indicated. See Table S9 for more details on species' traits. Altitude truncation was included in the GLMs only in response to difference in maximal elevation.

Extent	Response	SDM	Species' traits after stepwise selection	R^2
NSW	Current observed maximal altitude	-	Soil nutrients + Anthropochory	0.59
	Current predicted maximal altitude	Global	Subtropical	0.25
	Shift in current vs. future predicted maximal altitude	Global	Life strategies + elevation potential	0.83
	Current predicted maximal altitude	Local	Soil moisture + Life form + Zoochory + Subtropical	0.86
	Shift in current vs. future predicted maximal altitude	Local	Leaf duration	0.44
CH	Current observed maximal altitude	-	Temperature	0.53
	Current predicted maximal altitude	Global	Temperature + Soil nutrients	0.51
	Shift in current vs. future predicted maximal altitude	Global	elevation potential	0.32
	Current predicted maximal altitude	Local	Temperature + Rosette + Ploidy	0.70
	Shift in current vs. future predicted maximal altitude	Local	elevation potential	0.51

DISCUSSION

Our results show that, both local and global scales, SDM predict an average future increase of the altitudinal limit of invasive species in CH and NSW. Our study shows that montane, subalpine and even alpine levels in the most extreme climatic scenarios will be at higher risks of plant invasions in a future with warmer climates. However, the magnitude and the shape of this upward shift in response to global warming depend on the study area and the scale used for the calibration of SDM. In the following sections, we discuss the possible causes for these contrasted response patterns and we further present the implications of a distributional shift of invasive plant species at higher altitude for ecosystems and their conservation.

Why will plants move upward? The thermal vs. moisture hypothesis

The two study areas show different responses to global warming: in NSW, the shift is less severe than in CH. This is due to a different pool of species or different topo-climatic conditions between the study areas.

Compared to the pool of species in CH, species in NSW are adapted to colder environment, as shown by their lower Landolt's T values. This, in addition to warmer temperature at comparable elevation levels (Fig. S1), translates into species being currently distributed at a higher position among the vegetation levels in NSW, where species show an optimal suitability at the montane-subalpine levels whereas in CH, species are mostly potentially distributed at the lowland-montane levels. The selected

species in NSW are thus closer to the top of mountains than the selected species in CH, thus reducing their potential shift for elevation. It has been shown that current position along the elevation gradient is a major component explaining species elevational responses to climate warming (Engler *et al.*, 2011). Seen from a temperature gradient perspective, this means that a large part of the coldest conditions available in NSW are already suitable (and occupied) under current climatic conditions for most species, while they are not in CH (Fig. 7). As a result, suitable current areas are predicted to decrease more severely in NSW, as shown by Fig. 3.

However other factors may contribute to explain this different response between CH and NSW. For example, local SDM show that the alpine level is not currently highly suitable for most of species in NSW and won't become more suitable with climate warming. Here, different stressing factors shaping plant distribution under climate change in the two study areas may also explain this different pattern. In CH, invasive plant species are all clearly limited by cold in the local SDM whereas factors limiting species in NSW are less clearly defined (L1 on Fig. 2). In NSW, temperature seasonality and precipitation are significantly more important than in CH (L2 and L3 on Fig. 2).

To determine whether differences between CH and NSW are artifacts of species choice, we focused only on species present in both ranges, i.e. *Ailanthus altissima*, *Ambrosia artemisiifolia*, *Lonicera japonica, Prunus Laurocerasus* and *Robinia pseudoaccacia*. For these species global SDM calibrated on the worldwide extent show an upward shift in NSW (even more pronounced than in CH) whereas local SDM predict no shift when calibrated in NSW but still a continuous upward shift with warmer conditions when calibrated in CH (Fig. S7 and S8). The importance of climatic variables is also very different between local SDM calibrated in CH and NSW for identical species with a predominance of temperature variables in CH models and precipitation variables in NSW models (Fig. S9). Importantly, local SDMs in NSW tend to predict stress-tolerant species at lower maximal altitudes (negative coefficient for CSS species in Table S10), supporting the hypothesis that the main stress is aridity in NSW. The optimal water balance can shift downhill as shown for plants species in California and for vertebrates in different mountain chains among the world, thus reducing the upward component of the altitudinal shift (Crimmins *et al.*, 2011, McCain and Colwell, 2011).

The choice of the extent of the study areas, as it is known to have important implications for prediction accuracy (Barve *et al.*, 2011, Acevedo *et al.*, 2012, Rodriguez-Castaneda *et al.*, 2012). In our case however, study area extent is unlikely to have affected results because we chose to delimit SDM backgrounds for each species with the ecoregions where they were present. This approach has the advantage of delineating climatically similar for model building, although it may be too coarse for local fine scale SDM, especially if environmental gradients are particularly contrasted as precipitations in NSW.

Nevertheless, we can see strong overall convergences between our predictions under climate warming. In both CH and NSW cases, species are predicted to experience a decrease in the extent of their potential distribution as higher altitudinal levels become

rarer in the future in both CH and NSW landscapes (Fig. 3). What will happen to the lowland area is more uncertain. Global SDM also predict upward shifts of the lower species distributional limits but prediction for the lower distribution tails may be biased for two reasons. First, the database used to gather species distribution at global scale (GBIF) does not necessarily reflect total species distribution is known to be biased towards western European and northern American countries (Boakes *et al.*, 2010). Thus, we cannot exclude the possibility that our occurrence database for global SDM misses distribution in warmer areas of the world. Second, warmer areas distributed at the bottom of elevation gradients are the most affected by climate non-analogy for both global and local SDM (Fig. S2). If local SDM are extrapolated to these non-analog areas, which are highly unreliable because of spurious response curves, they show persistence at lower elevation in CH and no difference in NSW (Fig. S2). The PCA based on worldwide species distributions also shows that the CH environment should be more suitable in the future, in the contrast to NSW which shows a slight decrease of overall suitability. Thus, in addition to the entry of new invader species adapted to warmer conditions, persistence of the current pool of invaders at lower elevation cannot be excluded.

Patterns of invasions predicted by the models: which species are the mountain climbers?

It is known that lowland areas acts as a port of entry for invasive species and that only a subset of these can invade higher elevated areas (Pauchard *et al.*, 2009). However, to our knowledge, no study has yet identified which traits would enable invasive species to expand their altitudinal range.

In NSW, species preffering wetter soils occur at a higher elevation whereas species dispersed by anthropochory have a lower observed altitudinal limit (Table S10). In the Australian Alps, precipitation has a steeper gradient than temperature (Fig. S1). In this respect, the correlation between higher maximal observed elevation and higher precipitation is intuitive as is the negative association between anthropochory and altitude probably due to the decrease of anthropogenic activities in the wilder higher elevated areas (Pauchard *et al.*, 2009). Thus human brings fewer species than to lowlands. Predicted highest elevation is negatively related to the presence of species in subtropical areas. It seems reasonable that the few species of the NSW pool (N = 3) adapted to these hot and dry environments cannot reach the top of this wet and cold mountain chain. Additionally, species preferring higher soil moisture and those dispersed through zoochory are associated with a higher predicted altitude whereas woody life forms are more likely at lower predicted altitude (Table S10). This is in agreement with the sharp increase of precipitation at higher altitude in NSW (Fig. S1).

In CH, Landolt's T indicator is negatively correlated with actual and predicted elevation in both local and global SDM. This contrasted pattern with NSW shows that temperature is a more stressful factor in CH. On the contrary, soil moisture does not appear as a factor explaining species distribution along the altitudinal gradient. Additionally, the presence of

rosettes and diploidy is related to the ability of species to grow at higher predicted elevations in CH. Rosettes are typical alpine adaptations to light availability, nutrients storage, wind and herbivory regime (Billings, 1974, Choler, 2005, De Bello *et al.*, 2005). It is also common that diploid cytotypes are more adapted to higher elevation than tetraploids, which generally invade hotter areas (e.g. Gauthier *et al.*, 1998, Martin and Husband, 2013). More surprisingly, a preference for nutrient rich soils is also correlated with higher predicted altitude.

One important factor in the species distributional response to climate warming is the species elevation potential. Species with lower current predicted elevation, thus “further from the top”, show more predicted altitudinal expansion in both NSW and CH. Species have suitable habitats at higher altitudinal levels in NSW and many species show a future maximal predicted altitude corresponding the maximum the maximum possible altitude. However, the role of potential elevation is more intriguing in CH as species are less present at high altitudinal levels and only few reach the highest altitude of the gradient in the future (Table S1 and S2). Two hypotheses can explain this pattern. First, as species are more sensitive to temperature in CH, the magnitude of the response to climate warming is more important and their predicted altitudinal shift is so high that it theoretically reaches the nival limit at 3100 m. Second, the distributional response to global warming may not be linear along the altitudinal gradient affecting more severely lower than higher altitudes. Thus species at lower altitude would shift more than species already present at higher elevation. These two hypotheses are not exclusive. Additionally, the more deciduous the species, the higher species are predicted to shift in NSW with local SDM. Deciduous invasive species may be then more delimited by temperature in NSW than evergreen species.

Our traits results show that actual and predicted maximum elevation are limited by cold tolerance in CH and humidity tolerance in NSW. This difference supports our previous assertion that the main stressing factor shaping invasive species distribution along the elevational gradient are different in both study areas. Thus, the distribution response of species will depend on how the main distribution factors will be affected by climate change.

Implications for mid to high elevation ecosystems?

Currently, invasive plant species do not represent a major threat to biodiversity in CH because they are mostly distributed at lower elevation in human disturbed environments, which do not have high conservation value (Nobis *et al.*, 2009). If invasive plant species reach mountainous areas, they may have more severe effects than at lower elevation. Currently, montane and subalpine levels contain the highest species diversity in temperate climates like CH (Becker *et al.*, 2007). Our results suggest that these levels will be the most affected by the increase of suitability for invasive species. Although general theories predict that areas with higher species diversity will be more robust to perturbations of

invasive species (Elton, 1958, Tilman, 1999), supportive and non-supportive evidence exists (e.g. for: Kennedy *et al.*, 2002, Tilman *et al.*, 2006, against: Stohlgren *et al.*, 2003). Threats to mountain biodiversity from plant invasions can be even more severe because the current native biodiversity will be also affected by this upward distributional shift due to climate warming (Engler *et al.*, 2011). If species are naturally able to persist by leaving their current elevation to track their niche, they will rely on their dispersal abilities (Engler *et al.*, 2009). However, native species may also persist at their current elevation within micro-habitats (Scherrer and Korner, 2011). The additional rise of invasive species in these already disturbed communities can become more problematic for biodiversity. Moreover, mountains can be considered as island patches with more endemic, specialist, stress-tolerant and less competitive species, which are thus more vulnerable to biological invasions (Reaser *et al.*, 2007, Kueffer *et al.*, 2010). The trajectories of invasive species clearly need to be taken into account in planning biodiversity conservation in mountain regions (McDougall *et al.*, 2011*b*).

Limitations and perspectives

Our approach of stacking binarized predictions maps may overestimate the actual number of predicted species (Pottier *et al.*, 2013, Calabrese *et al.*, 2013, Dubuis *et al.*, *in rev.*) and should be rather considered as a suitability index for invasive plant species rather than the actual number of species. Macroecological models (MEM) depicting spatial distribution of species richness are known to be less biased but individual species responses to climate change are lost (Guisan and Rahbek, 2011). Previous MEM have been applied regionally in CH at a 1 km^2 resolution and provided a very similar picture of invasive plants “richness” under current conditions in CH (Nobis *et al.*, 2009). Additionally, the models may overestimate future potential distributions for two reasons. First, our approach did not include any soil or landuse variables and related change scenarios. Although these variables have limited impact for fine scale predictions in mountainous landscape (Randin *et al.*, 2009*b*), soil variables may be confounded with climatic variables, thus biasing and possibly overestimating future spatial predictions (Bertrand *et al.*, 2012). Second, our approach considers species dispersal ability as infinite. Dispersal limitation may explain why predicted highest elevation is higher than actual highest observed elevation for most of the species (Fig. 5, Table S1 to S4). While this unlimited dispersal scenario is over-simplistic, this approach is closer than the “no-dispersal” scenario in mountainous areas (Engler *et al.*, 2009). Moreover, invasive plant dispersal should not be limited along human and disturbed ways (i.e. roadsides, railways, rivers Pysek and Richardson, 2007). For natural or semi-natural habitats, the unlimited-dispersal assumption may not hold. On the other hand our approach may also underestimate future plant invaders suitability. Indeed, no demographical scenarios were applied to human disturbancs variable for future projections. In CH, demography and thus human disturbances is assumed to increase (Nobis *et al.*, 2009), which should favor invasive plant species. Our approach may be too conservative but as our and previous models (Nobis *et al.*, 2009) reveal, non-climatic

variables have much less relative importance in the depiction of potential distribution of invasive plant species.

Including soil variables, dispersal rates and more accurate scenarios for non-climatic variable is desirable, in particular for conservation plan requiring (very) fine scale modeling. However, this should not significantly impact our results and main conclusions.

CONCLUSION

Suitability for invasive species is not yet favorable at the highest vegetation levels in mountains. However, montane levels are already suitable in temperate mountain chains like NSW. Similarly as to the different degrees of invasions in the numerous mountain chains among the world, the response towards climate change could be more or less severe depending on the mountain chains locations. SDM predict an increase of suitability at higher altitudes, particularly in CH where species distribution is more limited by cold temperature. This represents a supplementary threat on mountainous biodiversity already affected by climate change. However, we have the rare advantage to prevent this additional threat to mountainous biodiversity. These evidences plead for a systematic consideration of invasive species in conservation plan concerning mountainous ecosystems.

Acknowledgments

Climatic scenarios data for local SDM in CH were provided by Achilleas Psomas, Dirk Schmarz and Nicklaus Zimmermann (WSL institute). Climatic scenario data for local SDM in NSW were provided by Christopher Simpson (NSW Office of Environment and Heritage). These latter layers were developed in an ARC Linkage project (LP0989537) on climate change, supported in part by GER. The computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics. AG, OB and BP received support at the earlier stages of the study from the Swiss National Center of Competences in Research ‘Plant Survival’ (granted by the Swiss National Science Foundation).

REFERENCES

Acevedo, P., Jiménez-Valverde, A., Lobo, J.M. & Real, R. (2012) Delimiting the geographical background in species distribution modelling. *Journal of Biogeography,* **39,** 1383-1390.

Alexander, J.M., Kueffer, C., Daehler, C.C., Edwards, P.J., Pauchard, A., Seipel, T. & Consortium, M. (2011) Assembly of nonnative floras along elevational gradients explained by directional ecological filtering. *Proceedings of the National Academy of Sciences,* **108,** 656-661.

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

Barni, E., Bacaro, G., Falzoi, S., Spanna, F. & Siniscalco, C. (2012) Establishing climatic constraints shaping the distribution of alien plant species along the elevation gradient in the Alps. *Plant Ecology,* **213,** 757-767.

Barton, K. (2012) MuMIn: Multi-model inference.

Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberon, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling,* **222,** 1810-1819.

Becker, A., Körner, C., Brun, J.-J., Guisan, A. & Tappeiner, U. (2007) Ecological and Land Use Studies Along Elevational Gradients. *Mountain Research and Development,* **27,** 58-65.

Becker, T., Dietz, H., Billeter, R., Buschmann, H. & Edwards, P.J. (2005) Altitudinal distribution of alien plant species in the Swiss Alps. *Perspectives in Plant Ecology Evolution and Systematics,* **7,** 173-183.

Benning, T.L., LaPointe, D., Atkinson, C.T. & Vitousek, P.M. (2002) Interactions of climate change with biological invasions and land use in the Hawaiian Islands: Modeling the fate of endemic birds using a geographic information system. *Proceedings of the National Academy of Sciences,* **99,** 14246-14249.

Bertrand, R., Perez, V. & Gégout, J.-C. (2012) Disregarding the edaphic dimension in species distribution models leads to the omission of crucial spatial information under climate change: the case of *Quercus pubescens* in France. *Global Change Biology,* **18,** 2648-2660.

Billings, W.D. (1974) Adaptations and Origins of Alpine Plants. *Arctic and Alpine Research,* **6,** 129-142.

Boakes, E.H., McGowan, P.J.K., Fuller, R.A., Chang-qing, D., Clark, N.E., O'Connor, K. & Mace, G.M. (2010) Distorted Views of Biodiversity: Spatial and Temporal Bias in Species Occurrence Data. *PLoS Biol,* **8,** e1000385.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters,* **4,** 585-589.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters,* **10,** 701-709.

Calabrese, J.M., Certain, G., Kraan, C. & Dormann, C.F. (2013) Stacking species distribution models and adjusting bias by linking them to macroecological models. *Global Ecology and Biogeography*.

Chefaoui, R.M. & Lobo, J.M. (2008) Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling,* **210,** 478-486.

Choler, P. (2005) Consistent Shifts in Alpine Plant Traits along a Mesotopographical Gradient. *Arctic, Antarctic, and Alpine Research,* **37,** 444-453.

Crimmins, S.M., Dobrowski, S.Z., Greenberg, J.A., Abatzoglou, J.T. & Mynsberge, A.R. (2011) Changes in Climatic Water Balance Drive Downhill Shifts in Plant Species' Optimum Elevations. *Science,* **331,** 324-327.

De Bello, F., Lepš, J.A.N. & Sebastià, M.-T. (2005) Predictive value of plant traits to grazing along a climatic gradient in the Mediterranean. *Journal of Applied Ecology,* **42,** 824-833.

Dray, S. & Dufour, A.B. (2007) The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software,* **22,** 1-20.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution,* **1,** 330-342.

Elton, C.S. (1958) *The Ecology of Invasions by Animals and Plants.* University of Chicago Press.

Engler, R., Randin, C.F., Thuiller, W., Dullinger, S., Zimmermann, N.E., Araujo, M.B., Pearman, P.B., Le Lay, G., Piedallu, C., Albert, C.H., Choler, P., Coldea, G., De Lamo, X., Dirnbock, T., Gegout, J.C., Gomez-Garcia, D., Grytnes, J.A., Heegaard, E., Hoistad, F., Nogues-Bravo, D., Normand, S., Puscas, M., Sebastia, M.T., Stanisci, A., Theurillat, J.P., Trivedi, M.R., Vittoz, P. & Guisan, A. (2011) 21st century climate change threatens mountain flora unequally across Europe. *Global Change Biology,* **17,** 2330-2341.

Engler, R., Randin, C.F., Vittoz, P., Czaka, T., Beniston, M., Zimmermann, N.E. & Guisan, A. (2009) Predicting future distributions of mountain plants under climate change: does dispersal capacity matter? *Ecography,* **32,** 34-45.

Fitzpatrick, M.C. & Hargrove, W.W. (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation,* **18,** 2255-2261.

Friedman, J.H., Hastie, T.J. & Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting. *Annals of Statistics,* **28,** 337-374.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology,* **98,** 790-799.

Gallien, L., Douzet, R., Pratte, S., Zimmermann, N.E. & Thuiller, W. (2012) Invasive species distribution models - how violating the equilibrium assumption can create new insights? *Global Ecology and Biogeography*.

Gauthier, P., Lumaret, R. & Bédécarrats, A. (1998) Genetic variation and gene flow in Alpine diploid and tetraploid populations of *Lotus* (*L. alpinus* (D.C.) Schleicher/*L. corniculatus* L.). I. Insights from morphological and allozyme markers. *Heredity,* **80,** 683-693.

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

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

Guisan, A. & Rahbek, C. (2011) SESAM - a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. *Journal of Biogeography,* **38,** 1433-1444.

Hay, S.I., Cox, J., Rogers, D.J., Randolph, S.E., Stern, D.I., Shanks, G.D., Myers, M.F. & Snow, R.W. (2002) Climate change and the resurgence of malaria in the East African highlands. *Nature,* **415,** 905-909.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology,* **25,** 1965-1978.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling,* **199,** 142-152.

Hufbauer, R.A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C.E., Rey, O. & Estoup, A. (2012) Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications,* **5,** 89-101.

Hulme, P.E. & Barrett, S.C.H. (2013) Integrating trait- and niche-based approaches to assess contemporary evolution in alien plant species. *Journal of Ecology,* **101,** 68-77.

IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, UK and New York, NY, USA.

Jakobs, G., Kueffer, C. & Daehler, C. (2010) Introduced weed richness across altitudinal gradients in Hawaii: humps, humans and water-energy dynamics. *Biological Invasions,* **12,** 4019-4031.

Juvik, J.O., Rodomsky, B.T., Price, J.P., Hansen, E.W. & Kueffer, C. (2011) The upper limits of vegetation on Mauna Loa, Hawaii: a 50th-anniversary reassessment. Ecology, 92, 518-525.

Kennedy, T.A., Naeem, S., Howe, K.M., Knops, J.M.H., Tilman, D. & Reich, P. (2002) Biodiversity as a barrier to ecological invasion. Nature, 417, 636-638.

Kueffer, C. (2010) Alien plants in the Alps: status and future invasion risks. Europe's ecological backbone: recognising the true value of our mountains (ed e. Price). European Environment Agency (EEA), Copenhagen.

Kueffer, C., Daehler, C.C., Torres-Santana, C.W., Lavergne, C., Meyer, J.-Y., Otto, R. & Silva, L. (2010) A global comparison of plant invasions on oceanic islands. Perspectives in Plant Ecology, Evolution and Systematics, 12, 145-161.

Kueffer, C., McDougall, K., Alexander, J., Daehler, C., Edwards, P.J., Haider, S., Milbau, A., Parks, C., Pauchard, A., Reshi, Z.A., Rew, L., Schroder, M. & Seipel, T. (2014) Plant invasions into mountain protected areas: assessment, prevention and control at multiple spatial scales. Springer, Dordrecht.

Kuhn, I., Durka, W. & Klotz, S. (2004) BiolFlor - a new plant-trait database as a tool for plant invasion ecology. Diversity and Distributions, 10, 363-365.

Kurz, W.A., Dymond, C.C., Stinson, G., Rampley, G.J., Neilson, E.T., Carroll, A.L., Ebata, T. & Safranyik, L. (2008) Mountain pine beetle and forest carbon feedback to climate change. Nature, 452, 987-990.

Landolt, E. (2010) *Flora indicativa: Ecological indicator values and Biological attributes of flora of Switzerland and the Alps.* Haupt, Bern.

Leung, B., Lodge, D.M., Finnoff, D., Shogren, J.F., Lewis, M.A. & Lamberti, G. (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **269,** 2407-2413.

Mandle, L., Warren, D.L., Hoffmann, M.H., Peterson, A.T., Schmitt, J. & von Wettberg, E.J. (2010) Conclusions about Niche Expansion in Introduced *Impatiens walleriana* Populations Depend on Method of Analysis. *Plos One,* **5**.

Marini, L., Gaston, K.J., Prosser, F. & Hulme, P.E. (2009) Contrasting response of native and alien plant species richness to environmental energy and human impact along alpine elevation gradients. *Global Ecology and Biogeography,* **18,** 652-661.

Martin, S.L. & Husband, B.C. (2013) Adaptation of diploid and tetraploid *Chamerion angustifolium* to elevation but not local environment. *Evolution,* **67,** 1780-1791.

McCain, C.M. (2007) Could temperature and water availability drive elevational species richness patterns? A global case study for bats. *Global Ecology and Biogeography,* **16,** 1-13.

McCain, C.M. & Colwell, R.K. (2011) Assessing the threat to montane biodiversity from discordant shifts in temperature and precipitation in a changing climate. *Ecology Letters,* **14,** 1236-1245.

McCullagh, P. & Nelder, J.A. (1983) *Generalized Linear Models: First Edition.* Chapman and Hall, London, New-York.

McDougall, K.L., Alexander, J.M., Haider, S., Pauchard, A., Walsh, N.G. & Kueffer, C. (2011*a*) Alien flora of mountains: global comparisons for the development of local preventive measures against plant invasions. *Diversity and Distributions,* **17,** 103-111.

McDougall, K.L., Khuroo, A.A., Loope, L.L., Parks, C.G., Pauchard, A., Reshi, Z.A., Rushworth, I. & Kueffer, C. (2011*b*) Plant Invasions in Mountains: Global Lessons for Better Management. *Mountain Research and Development,* **31,** 380-387.

McDougall, K.L., Morgan, J.W., Walsh, N.G. & Williams, R.J. (2005) Plant invasions in treeless vegetation of the Australian Alps. *Perspectives in Plant Ecology, Evolution and Systematics,* **7,** 159-171.

Moles, A.T., Gruber, M.A.M. & Bonser, S.P. (2008) A new framework for predicting invasive plant species. *Journal of Ecology,* **96,** 13-17.

Nobis, M.P., Jaeger, J.A.G. & Zimmermann, N.E. (2009) Neophyte species richness at the landscape scale under urban sprawl and climate warming. *Diversity and Distributions,* **15,** 928-939.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. (2001) Terrestrial ecoregions of the worlds: A new map of life on Earth. *Bioscience,* **51,** 933-938.

Parks, C.G., Radosevich, S.R., Endress, B.A., Naylor, B.J., Anzinger, D., Rew, L.J., Maxwell, B.D. & Dwire, K.A. (2005) Natural and land-use history of the Northwest mountain ecoregions (USA) in relation to patterns of plant invasions. *Perspectives in Plant Ecology, Evolution and Systematics,* **7,** 137-158.

Pauchard, A. & Alaback, P.B. (2004) Influence of Elevation, Land Use, and Landscape Context on Patterns of Alien Plant Invasions along Roadsides in Protected Areas of South-Central Chile. *Conservation Biology,* **18,** 238-248.

Pauchard, A., Kueffer, C., Dietz, H., Daehler, C.C., Alexander, J., Edwards, P.J., Arévalo, J.R., Billeter, R., Cavieres, L., Guisan, A., Haider, S., Jakobs, G., Mack, R.N., McDougall, K., Millar, C., Naylor, B.J., Parks, C., Poll, M., Rew, L.J. & Seipel, T. (2009) Ain't no mountain high enough: Plant invasions reaching high elevations. *Frontiers in Ecology and the Environment,* **7,** 479-486.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography,* **38,** 817-827.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science,* **335,** 1344-1348.

Petitpierre, B., Broennimann, O., Kueffer C., Daehler C., Guisan A. (**In Prep.**) The importance of variable selection in species distribution models: the case of biological invasions.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* **190,** 231-259.

Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C.F., Vittoz, P. & Guisan, A. (2013) The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Global Ecology and Biogeography,* **22,** 52-63.

Pysek, P. & Richardson, D.M. (2007) Traits associated with invasiveness in alien plants: where do we stand? *Biological Invasions*, pp. 97-125. Springer.

R (2005) *R : A language and environment for statistical computing. Version 2.5.1.* Development Core Team. R Foundation for Statistical Computing, Vienna, Austria.

Rahbek, C. (1995) The elevational gradient of species richness: a uniform pattern? *Ecography,* **18,** 200-205.

Ramirez, J. & Jarvis, A. (2008) High Resolution Statistically Downscaled Future Climate Surfaces. International Center for Tropical Agriculture (CIAT). CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). Cali, Colombia.

Randin, C.F., Engler, R., Normand, S., Zappa, M., Zimmermann, N.E., Pearman, P.B., Vittoz, P., Thuiller, W. & Guisan, A. (2009*a*) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology,* **15,** 1557-1569.

Randin, C.F., Vittoz, P., Jaccard, H., Yoccoz, N.G. & Guisan, A. (2009*b*) Land use improves spatial predictions of plant abundances but not occurrences in a mountain landscape. *Journal of Vegetation Science,* **20,** 996-1008.

Reaser, J.K., Meyerson, L.A., Cronk, Q., De Poorter, M., Eldrege, L., Green, E., Kairo, M., Latasi, P., Mack, R.N. & Mauremootoo, J. (2007) Ecological and socioeconomic impacts of invasive alien species in island ecosystems. *Environmental Conservation,* **34,** 98-111.

Rodriguez-Castaneda, G., Hof, A.R., Jansson, R. & Harding, L.E. (2012) Predicting the Fate of Biodiversity Using Species' Distribution Models: Enhancing Model Comparability and Repeatability. *Plos One,* **7,** e44402.

Romdal, T.S. & Grytnes, J.-A. (2007) An indirect area effect on elevational species richness patterns. *Ecography,* **30,** 440-448.

Scherrer, D. & Korner, C. (2011) Topographically controlled thermal-habitat differentiation buffers alpine plant diversity against climate warming. *Journal of Biogeography,* **38,** 406-416.

Seipel, T. (2011) *Distribution and demographics of non-native plant species in mountainous regions. Diss Nr. 20031*. PhD, ETH-Zürich, Zürich.

Seipel, T., Alexander, J.M., Daehler, C.C., Edwards, P.J., Dar, P.A., McDougall, K., Naylor, N., Parks, C., Reshi, Z.A., J., R.L., M., S. & C., K. (in revision) Home and away: Performance of an invasive plant in native and non-native regions. *Journal of Biogeography*.

Seipel, T., Kueffer, C., Rew, L.J., Daehler, C.C., Pauchard, A., Naylor, B.J., Alexander, J.M., Edwards, P.J., Parks, C.G., Arevalo, J.R., Cavieres, L.A., Dietz, H., Jakobs, G., McDougall, K., Otto, R. & Walsh, N. (2012) Processes at multiple scales affect richness and similarity of non-native plant species in mountains around the world. *Global Ecology and Biogeography,* **21,** 236-246.

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

Stohlgren, T.J., Barnett, D.T. & Kartesz, J.T. (2003) The rich get richer: patterns of plant invasions in the United States. *Frontiers in Ecology and the Environment,* **1,** 11-14.

Thompson, K. & Davis, M.A. (2011) Why research on traits of invasive plants tells us very little. *Trends in ecology & evolution (Personal edition),* **26,** 155-156.

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

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

Tilman, D. (1999) The ecological consequences of changes in biodiversity: a search for general principles. *Ecology,* **80,** 1455-1474.

Tilman, D., Reich, P.B. & Knops, J.M.H. (2006) Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature,* **441,** 629-632.

Trtikova, M., Edwards, P.J. & Güsewell, S. (2010) No adaptation to altitude in the invasive plant *Erigeron annuus* in the Swiss Alps. *Ecography,* **33,** 556-564.

Van Kleunen, M., Weber, E. & Fischer, M. (2010) A meta-analysis of trait differences between invasive and non-invasive plant species. *Ecology Letters,* **13,** 235-245.

Walther, G.-R., Roques, A., Hulme, P.E., Sykes, M.T., Pyšek, P., Kühn, I., Zobel, M., Bacher, S., Botta-Dukát, Z., Bugmann, H., Czúcz, B.l., Dauber, J., Hickler, T., Jarošik, V.c., Kenis, M., Klotz, S., Minchin, D., Moora, M., Nentwig, W., Ott, J.r., Panov, V.E., Reineking, B.r., Robinet, C., Semenchenko, V., Solarz, W., Thuiller, W., VilÃ , M., Vohland, K. & Settele, J. (2009) Alien species in a warmer world: risks and opportunities. *Trends in Ecology & Evolution,* **24,** 686-693.

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

Zomer, R.J., Trabucco, A., Bossio, D.A., van Straaten, O. & Verchot, L.V. (2008) Climate Change Mitigation: A Spatial Analysis of Global Land Suitability for Clean Development Mechanism Afforestation and Reforestation. *Agriculture, Ecosystems & Environment,* **126,** 67 - 80.

Zweig, M.H. & Campbell, G. (1993) Receiver operating characteristic (roc) plots - a fundamental evaluation tool in clinical medicine (vol 39, pg 561, 1993). *Clinical Chemistry,* **39,** 1589-1589.

SUPPLEMENTARY MATERIALS

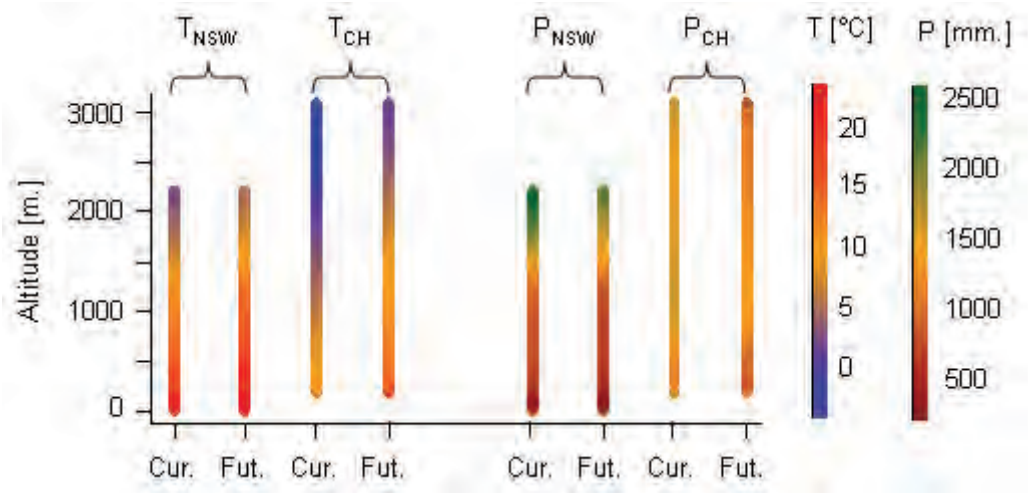


Figure S1: Average annual temperature (T) and annual precipitation (P) under current (Cur.) and future (Hadcm3 A1 2070, Fut.) conditions in New South Wales (NSW) and Switzerland (CH) along altitudinal gradient.

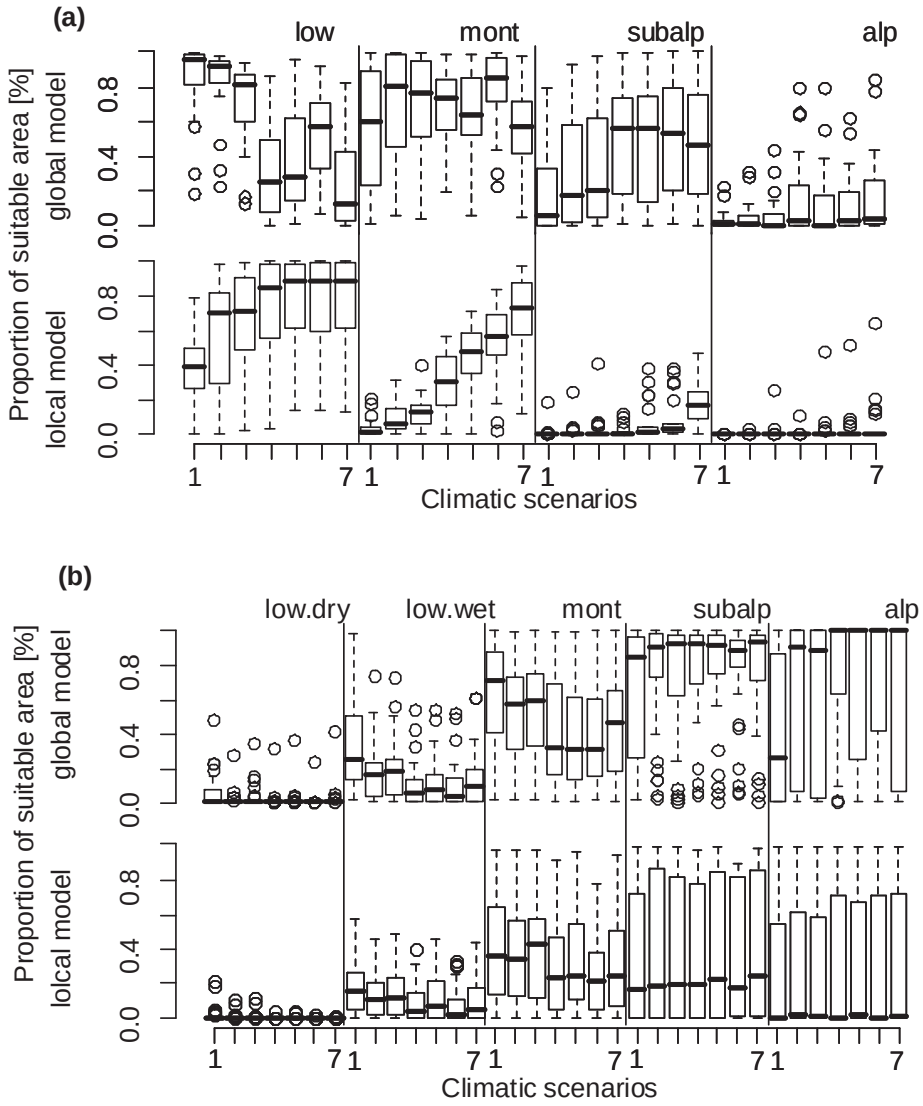


Figure S2 Proportion of suitable area predicted by SDM for the seven climatic scenarios (ranked from 1 to 7, follwowing Table S5 - S8) at each vegetation level (lowland, montane, subalpine and alpine) in Switzerland (a) and New South Wales (b) without considering future climate non-analogy.

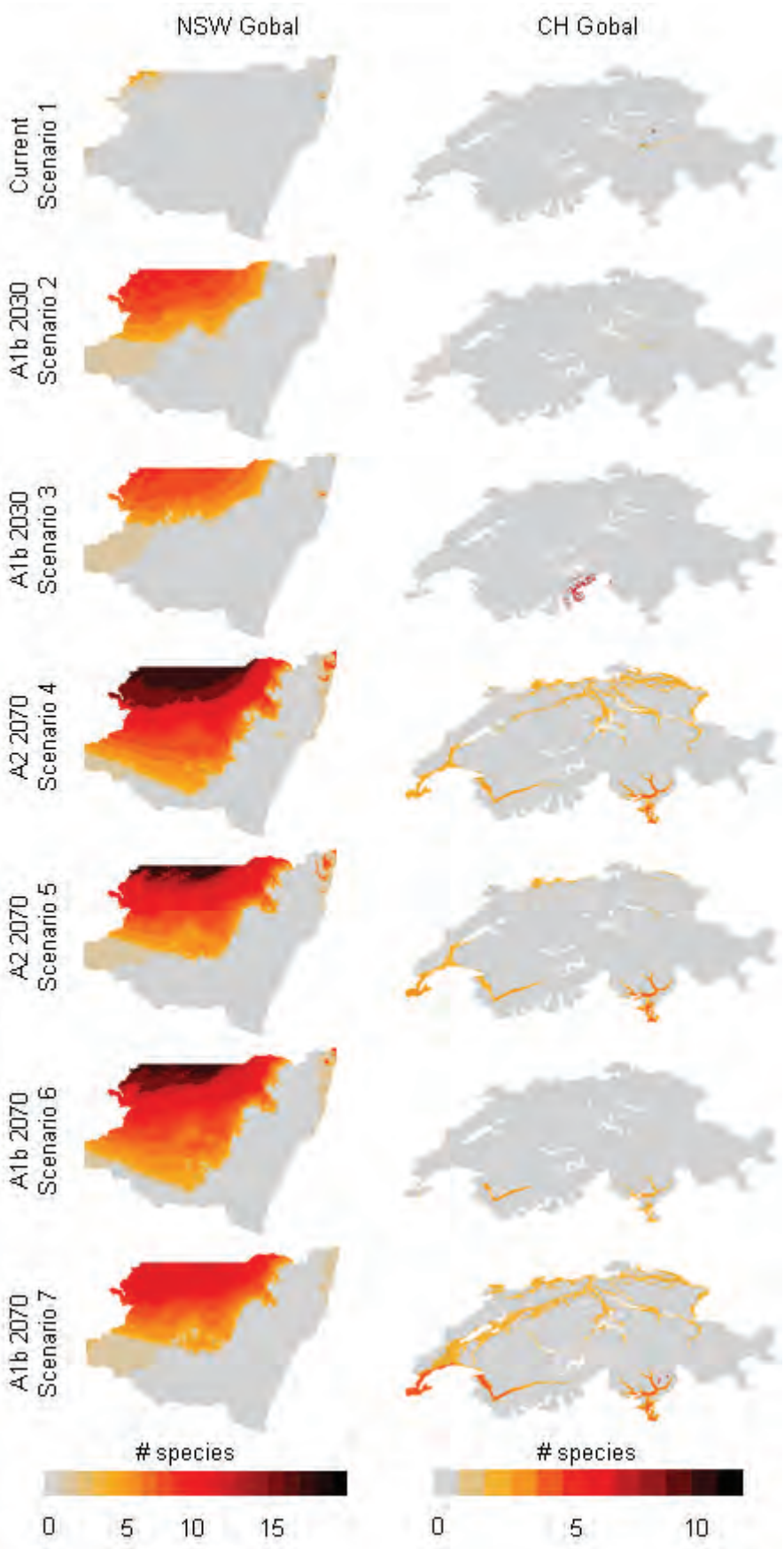


Figure S3: Non-analog climates in New South Wales and Switzerland under current and future climatic scenarios. For each species, sites of the study area were classified as analog (0) or non-analog (1) based on the MESS value calculated with their global niche. These layers were then piled to show the number of species for which climate is non-analog.

https://drive.google.com/file/d/0B5ZZgCyoOEdnWDlaVmNSQmdPNVk/edit?usp=sharing

Figure S4: Global SDM results for each species. Extent and occurrences used to calibrate SDM are shown as grey and red dots respectively. Barplots represent relative variables importance and boxplots represent continuous suitability for each vegetation level. Continuous suitability are spatially represented under current and future climate scenario, so is MESS values. Black dots represent actual species distribution in the study area**.**

https://drive.google.com/file/d/0B5ZZgCyoOEdnTHRYNEdENlJid1E/edit?usp=sharing

Figure S5: Local SDM results for each species. Barplots represent relative variables importance and boxplots represent continuous suitability for each vegetation level. Continuous suitability are spatially represented under current and future climate scenario. Black dots represent actual species distribution in the study area.

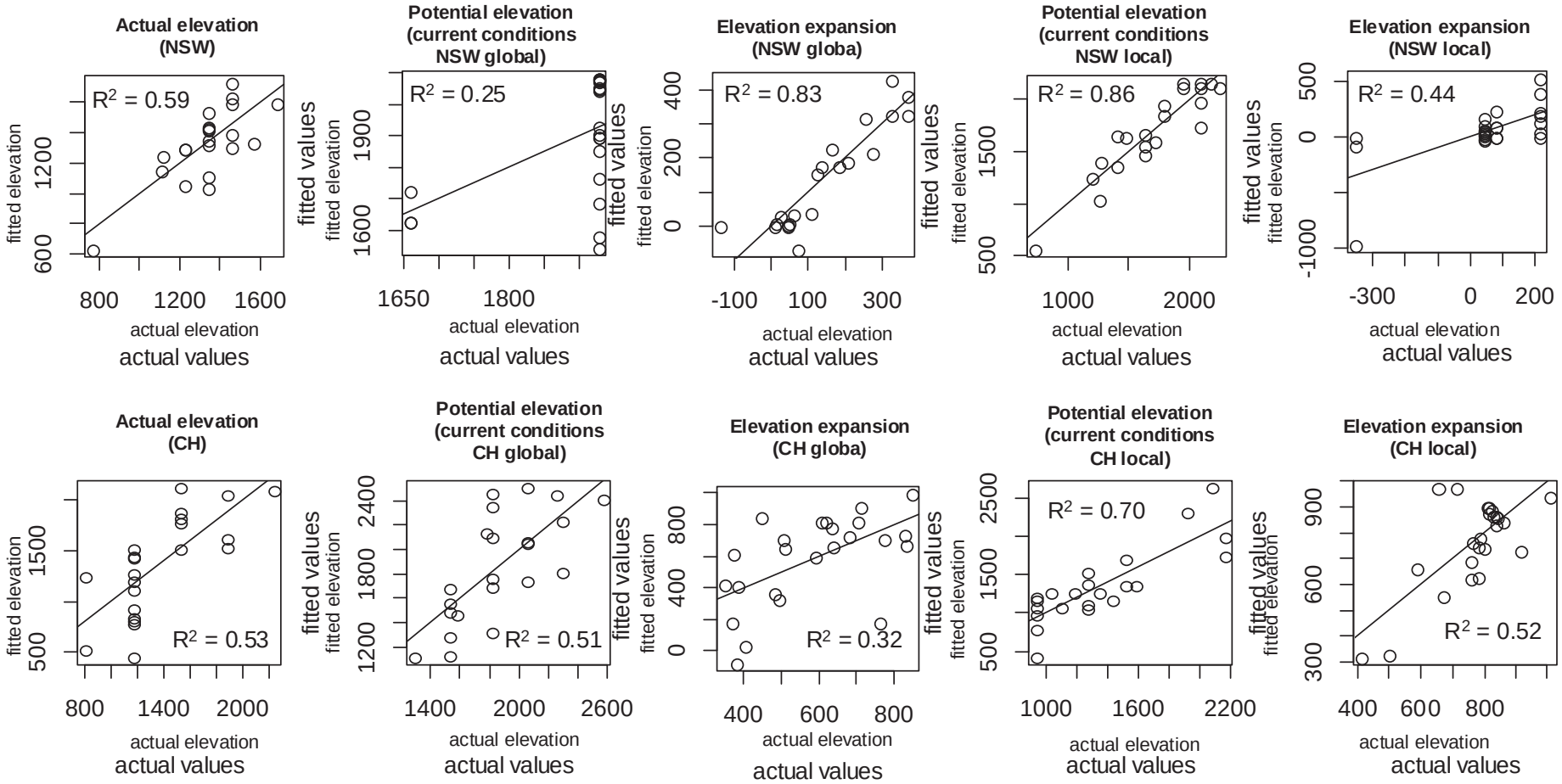


Figure S6: Models calibrated on traits to explain actual highest altitude, highest predicted altitude and predicted altitudinal expansion of species in Switzerland (CH) and New South Wales (NSW).

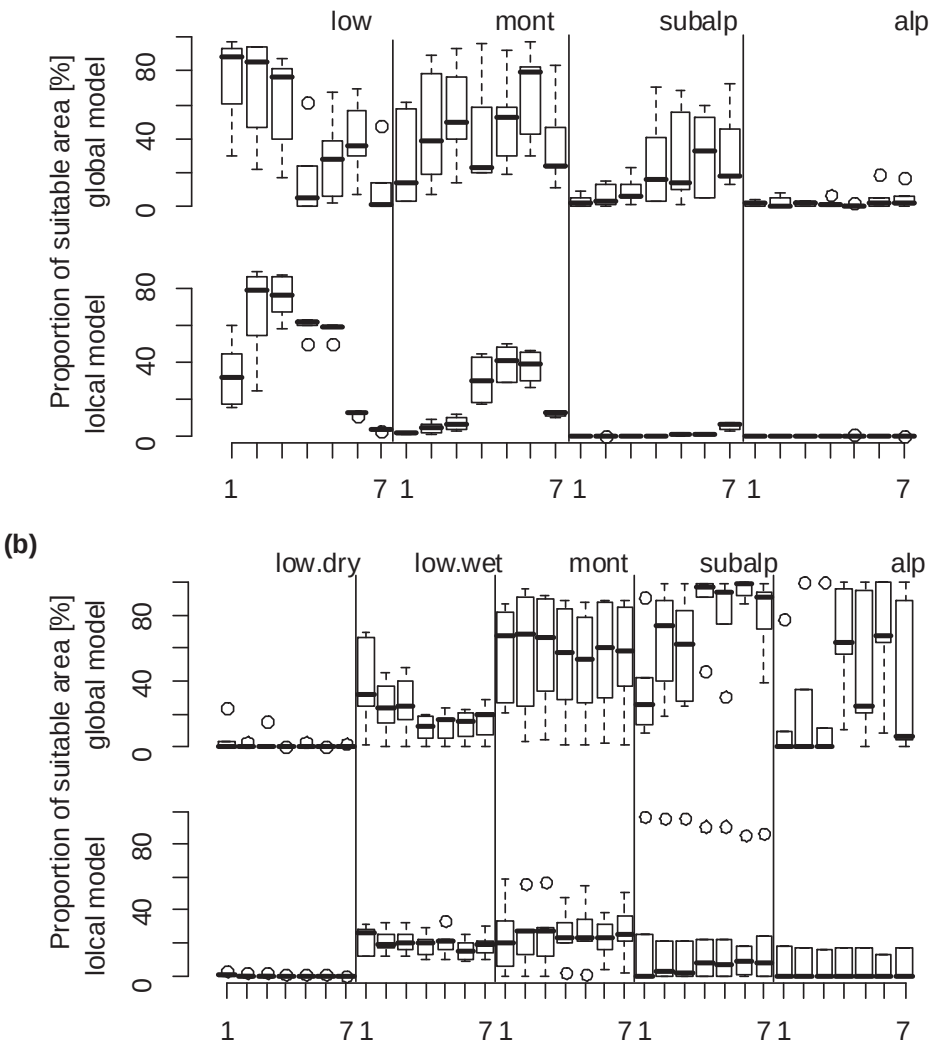


Figure S7: Proportion of suitable area predicted by global and local SDM for the five species present in both Switzerland (a) and New South Wales (b), i. e. *Ailanthus altissima*, *Ambrosia artemisiifolia*, *Lonicera japonica, Prunus Laurocerasus* and *Robinia pseudoaccacia*. Suitable areas is assessed for the seven climatic scenarios (ranked from 1 to 7, following Table S5 to S8) at each vegetation level (lowland, montane, subalpine and alpine)

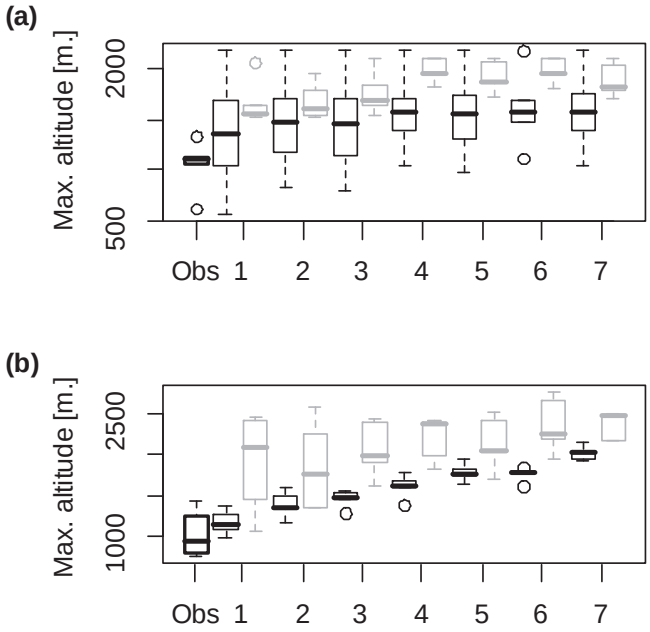


Figure S8: Highest observed (Obs) and predicted altitude following different climatic scenarios (ranked from 1 to 7) predicted by local (black) and global (grey) models in Switzerland (a) and New South Wales (b) for the five species present in both Switzerland and New South Wales, i. e. *Ailanthus altissima*, *Ambrosia artemisiifolia*, *Lonicera japonica, Prunus Laurocerasus* and *Robinia pseudoaccacia*.

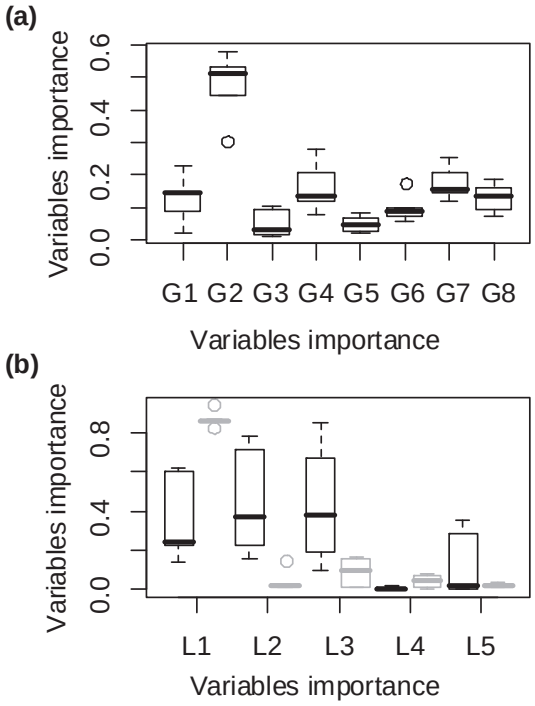


Figure S9: Variables importance (abbreviated as in Table 1) for the five species present in both Switzerland and New South Wales used in global (a) and local SDM (b). At the local scale, variable importance for SDM in NSW and CH is drawn in black and grey respectively.

Table S1: Results for CH global models. This table reports number of occurrences used in the models (N_{occ}), evaluation of models for each species (AUC, TSS, Boyce, Sensitivity with best TSS threshold), importance of variables (G1-G8) and actual observed elevation (E_a), predicted elevation under current climate (E_p) and future climate (E_{2070}, scenario: A1B; GCM: averaged HadCM3 and Echam5)

Species	Genus	N_{occ}	AUC	TSS	Boyce	S_{TSS}	G1	G2	G3	G4	G5	G6	G7	G8	E_a	E_p	E_{2070}
Ailanthus	*altissima*	175	0.82	0.53	0.92	0.72	0.15	0.30	0.03	0.21	0.08	0.07	0.25	0.16	800	2093	2418
Ambrosia	*artemisiifolia*	259	0.86	0.57	0.96	0.74	0.02	0.51	0.09	0.12	0.02	0.18	0.14	0.09	1242	1458	2368
Amorpha	*fruticosa*	89	0.83	0.63	0.79	0.78	0.01	0.22	0.01	0.37	0.20	0.06	0.08	0.20	446	1317	1501
Artemisia	*verlotiorum*	143	0.84	0.57	0.94	0.79	0.03	0.18	0.39	0.15	0.33	0.12	0.28	0.25	1520	1691	2473
Buddleja	*davidii*	334	0.92	0.72	0.95	0.86	0.25	0.10	0.01	0.25	0.03	0.13	0.27	0.11	1196	1482	2299
Bunias	*orientalis*	143	0.84	0.53	0.90	0.84	0.07	0.45	0.07	0.26	0.02	0.01	0.29	0.17	1611	2232	3080
Cornus	*sericea*	141	0.82	0.54	0.76	0.81	0.07	0.48	0.01	0.19	0.07	0.18	0.10	0.04	925	1282	1993
Cyperus	*esculentus*	128	0.80	0.49	0.85	0.69	0.47	0.08	0.15	0.08	0.14	0.04	0.16	0.14	512	1114	1782
Erigeron	*annuus*	212	0.94	0.77	0.94	0.87	0.03	0.60	0.02	0.11	0.02	0.22	0.17	0.12	1875	2506	2925
Helianthus	*tuberosus*	108	0.87	0.61	0.88	0.76	0.03	0.59	0.01	0.28	0.02	0.05	0.19	0.09	1264	1762	2581
Heracleum	*mantegazzianum*	416	0.93	0.73	0.92	0.88	0.07	0.37	0.01	0.26	0.01	0.08	0.21	0.03	2051	2407	2818
Impatiens	*glandulifera*	632	0.93	0.75	0.97	0.89	0.09	0.32	0.12	0.09	0.02	0.04	0.35	0.06	1520	1735	2556
Lonicera	*japonica*	207	0.85	0.55	0.96	0.79	0.23	0.45	0.10	0.28	0.07	0.10	0.15	0.14	932	2420	2341
Lupinus	*polyphyllus*	357	0.89	0.65	0.96	0.88	0.08	0.53	0.11	0.05	0.05	0.04	0.26	0.14	2088	2439	3054
Mahonia	*aquifolium*	109	0.92	0.70	0.82	0.83	0.16	0.33	0.01	0.23	0.07	0.06	0.11	0.03	838	1127	1859
Phytolacca	*americana*	115	0.85	0.59	0.80	0.82	0.45	0.17	0.30	0.03	0.07	0.02	0.22	0.25	778	2350	2378
Prunus	*laurocerasus*	240	0.93	0.78	0.95	0.89	0.14	0.53	0.01	0.14	0.04	0.09	0.12	0.07	764	1067	2062
Reynoutria	*japonica*	137	0.94	0.81	0.90	0.90	0.03	0.38	0.02	0.12	0.10	0.59	0.19	0.13	1825	2063	2769
Rhus	*typhina*	72	0.92	0.73	0.79	0.81	0.13	0.05	0.03	0.32	0.07	0.29	0.16	0.32	1111	1681	2344
Robinia	*pseudoacacia*	472	0.90	0.63	0.94	0.77	0.09	0.58	0.02	0.08	0.02	0.05	0.20	0.19	1433	2458	2640
Sedum	*spurium*	92	0.91	0.71	0.81	0.85	0.03	0.43	0.02	0.17	0.01	0.09	0.23	0.04	1775	2131	2493
Senecio	*inaequidens*	174	0.94	0.79	0.81	0.88	0.07	0.37	0.01	0.13	0.04	0.15	0.17	0.13	1448	1553	2279
Solidago	*canadensis*	431	0.91	0.68	0.92	0.80	0.04	0.33	0.02	0.14	0.01	0.04	0.26	0.05	2124	2050	2701
Solidago	*gigantea*	276	0.95	0.77	0.91	0.86	0.07	0.38	0.01	0.23	0.01	0.05	0.28	0.04	1523	1810	2404

Table S2: Results for CH local models. This table reports number of occurrences used in the models (N_{occ}), evaluation of models for each species (AUC, TSS, Boyce, Sensitivity with best TSS threshold), importance of variables (L1-L5) and actual observed elevation (E_a), predicted elevation under current climate (E_p) and future climate (E_{2070}, scenario: A1B; GCM: averaged HadCM3 and Echam5)

Genus	Species	N_{occ}	AUC	TSS	Boyce	S_{TSS}	L1	L2	L3	L4	L5	E_a	E_p	E_{2070}
Ailanthus	*altissima*	93	0.98	0.89	0.83	0.96	0.86	0.02	0.16	0.08	0.01	800	980	1777
Ambrosia	*artemisiifolia*	214	0.97	0.86	0.92	0.93	0.87	0.01	0.01	0.08	0.01	1242	1267	1862
Amorpha	*fruticosa*	11	1.00	1.00	0.40	1.00	0.16	0.01	0.83	0.00	0.12	446	425	1331
Artemisia	*verlotiorum*	78	0.90	0.69	0.90	0.93	0.99	0.02	0.06	0.03	0.04	1520	1249	1976
Buddleja	*davidii*	310	0.88	0.61	0.91	0.87	0.88	0.01	0.07	0.01	0.03	1196	1154	2016
Bunias	*orientalis*	68	0.91	0.74	0.91	0.94	0.55	0.57	0.11	0.00	0.02	1611	1737	2698
Cornus	*sericea*	11	0.95	0.88	0.45	1.00	0.48	0.26	0.21	0.02	0.48	925	1527	2452
Cyperus	*esculentus*	13	0.93	0.86	0.25	1.00	0.93	0.10	0.21	0.00	0.04	512	1066	1880
Erigeron	*annuus*	370	0.80	0.46	0.92	0.83	0.92	0.02	0.02	0.02	0.01	1875	1363	2024
Helianthus	*tuberosus*	48	0.92	0.76	0.69	0.93	0.91	0.03	0.06	0.04	0.13	1264	1205	1898
Heracleum	*mantegazzianum*	321	0.78	0.43	0.95	0.74	0.67	0.03	0.04	0.04	0.15	2051	1990	2632
Impatiens	*glandulifera*	278	0.90	0.65	0.92	0.87	0.92	0.03	0.02	0.00	0.26	1520	1347	2101
Lonicera	*japonica*	20	0.96	0.86	0.21	1.00	0.95	0.15	0.10	0.00	0.02	932	1090	1908
Lupinus	*polyphyllus*	50	0.94	0.81	0.84	0.93	0.36	0.07	0.04	0.34	0.46	2088	2639	2910
Mahonia	*aquifolium*	38	0.94	0.79	0.85	0.93	0.83	0.02	0.03	0.15	0.07	838	1050	1835
Phytolacca	*americana*	25	1.00	0.99	0.54	1.00	0.99	0.11	0.05	0.02	0.03	778	773	1484
Prunus	*laurocerasus*	59	0.91	0.69	0.89	0.96	0.83	0.02	0.16	0.01	0.03	764	1146	1975
Reynoutria	*japonica*	389	0.89	0.62	0.91	0.83	0.91	0.01	0.01	0.02	0.06	1825	1262	1982
Rhus	*typhina*	105	0.91	0.71	0.86	0.89	0.95	0.01	0.02	0.02	0.09	1111	1115	1945
Robinia	*pseudoacacia*	251	0.89	0.62	0.95	0.82	0.86	0.03	0.01	0.05	0.04	1433	1366	1944
Sedum	*spurium*	40	0.84	0.61	0.67	0.84	0.42	0.07	0.02	0.39	0.06	1775	2309	2561
Senecio	*inaequidens*	115	0.93	0.76	0.77	0.87	0.95	0.05	0.02	0.01	0.01	1448	1064	1845
Solidago	*canadensis*	393	0.87	0.59	0.92	0.85	0.67	0.09	0.02	0.04	0.03	2124	1690	2226
Solidago	*gigantea*	303	0.89	0.68	0.84	0.84	0.75	0.01	0.10	0.02	0.11	1523	1164	2031

Table S3: Results for NSW global models. This table reports number of occurrences used in the models (Nocc), evaluation of models for each species (AUC, TSS, Boyce, Sensitivity with best TSS threshold), importance of variables (G1-G8) and actual observed elevation (Ea), predicted elevation under current climate (Ep) and future climate (E2070, scenario: A1B; GCM: HadCM3 and Echam5).

Genus	species	N_{occ}	AUC	TSS	Boyce	S_{TSS}	G1	G2	G3	G4	G5	G6	G7	G8	E_a	E_p	E_{2070}
Ageratina	*adenophora*	164	0.95	0.83	0.83	0.93	0.06	0.24	0.14	0.00	0.02	0.33	0.15	0.11	1034	1262	1498
Ailanthus	*altissima*	175	0.82	0.53	0.92	0.72	0.15	0.30	0.03	0.21	0.08	0.07	0.25	0.16	1056	1543	1871
Ambrosia	*artemisiifolia*	259	0.86	0.57	0.96	0.74	0.02	0.51	0.09	0.12	0.02	0.18	0.14	0.09	632	1628	2060
Anthoxanthum	*odoratum*	3868	0.90	0.68	0.99	0.88	0.35	0.74	0.21	0.17	0.08	0.07	0.11	0.06	1726	2083	2083
Carduus	*nutans*	269	0.91	0.69	0.86	0.81	0.12	0.53	0.06	0.26	0.04	0.03	0.09	0.22	1289	1897	2074
Chondrilla	*juncea*	1071	0.90	0.69	0.98	0.91	0.37	0.48	0.33	0.16	0.05	0.08	0.31	0.35	1147	1685	1615
Cirsium	*vulgare*	3190	0.91	0.68	0.99	0.90	0.20	0.56	0.06	0.42	0.03	0.12	0.10	0.07	1629	1764	2083
Conium	*maculatum*	728	0.89	0.68	0.95	0.90	0.29	0.53	0.11	0.37	0.05	0.12	0.14	0.03	1244	1724	1938
Convolvulus	*arvensis*	2501	0.94	0.76	0.97	0.88	0.29	0.56	0.26	0.35	0.06	0.13	0.14	0.05	1320	1628	1954
Cortaderia	*selloana*	100	0.82	0.57	0.79	0.71	0.07	0.46	0.03	0.16	0.04	0.11	0.01	0.39	303	1090	1365
Crataegus	*monogyna*	3095	0.91	0.69	0.97	0.93	0.23	0.48	0.01	0.50	0.08	0.09	0.09	0.04	1335	1670	2018
Cytisus	*scoparius*	798	0.91	0.68	0.96	0.84	0.04	0.06	0.04	0.63	0.06	0.10	0.01	0.06	1528	1582	1968
Echium	*vulgare*	771	0.91	0.69	0.96	0.87	0.17	0.69	0.04	0.11	0.06	0.12	0.13	0.18	1297	2046	2083
Eragrostis	*curvula*	314	0.87	0.61	0.94	0.79	0.06	0.41	0.01	0.02	0.12	0.13	0.55	0.13	1333	1541	1593
Hedera	*helix*	2139	0.89	0.64	0.99	0.84	0.17	0.70	0.10	0.10	0.09	0.04	0.06	0.07	1298	1894	2083
Holcus	*lanatus*	3033	0.91	0.70	0.98	0.90	0.27	0.51	0.04	0.45	0.08	0.08	0.11	0.02	1594	1856	2083
Hypericum	*perforatum*	2030	0.91	0.70	0.97	0.90	0.35	0.79	0.25	0.04	0.07	0.07	0.14	0.04	1390	1927	2083
Ilex	*aquifolium*	1474	0.94	0.78	0.98	0.92	0.46	0.43	0.03	0.30	0.14	0.08	0.34	0.05	1036	2075	2083
Leucanthemum	*vulgare*	2075	0.91	0.71	0.99	0.89	0.34	0.68	0.19	0.21	0.06	0.08	0.19	0.09	1417	1906	2083
Lonicera	*japonica*	207	0.85	0.55	0.96	0.79	0.23	0.45	0.10	0.28	0.07	0.10	0.15	0.14	1332	1526	1738
Lotus	*corniculatus*	3370	0.93	0.75	0.98	0.90	0.40	0.61	0.21	0.21	0.16	0.06	0.17	0.05	1350	2076	2083
Mimulus	*moschatus*	99	0.84	0.57	0.87	0.87	0.01	0.74	0.04	0.01	0.18	0.04	0.11	0.07	1591	2083	2083
Pinus	*radiata*	130	0.85	0.58	0.89	0.85	0.05	0.19	0.38	0.05	0.02	0.05	0.09	0.23	1576	1530	1705
Poa	*pratensis*	3747	0.89	0.63	0.99	0.80	0.25	0.56	0.12	0.13	0.10	0.05	0.09	0.09	1336	2078	2083
Prunus	*laurocerasus*	240	0.93	0.78	0.95	0.89	0.14	0.53	0.01	0.14	0.04	0.09	0.12	0.07	1132	1559	1848
Robinia	*pseudoacacia*	472	0.90	0.63	0.94	0.77	0.09	0.58	0.02	0.08	0.02	0.05	0.20	0.19	1110	2054	2083
Rosa	*rubiginosa*	667	0.93	0.73	0.97	0.85	0.03	0.88	0.01	0.02	0.01	0.06	0.02	0.17	1429	2083	2083
Tragopogon	*dubius*	252	0.86	0.61	0.94	0.73	0.01	0.52	0.02	0.08	0.02	0.26	0.21	0.45	1416	2071	2051
Verbascum	*thapsus*	989	0.88	0.64	0.97	0.85	0.12	0.60	0.02	0.02	0.08	0.09	0.29	0.14	1440	2051	2083

Table S4: Results for NSW local. This table reports number of occurrences used in the models (N_{occ}), evaluation of models for each species (AUC, TSS, Boyce, Sensitivity with best TSS threshold), importance of variables (L1-L5) and actual observed elevation (E_a), predicted elevation under current climate (E_p) and future climate (E_{2070}, scenario: A1B; GCM: HadCM3 and Echam5)

Genus	species	N_{occ}	AUC	TSS	Boyce	S_{TSS}	L1	L2	L3	L4	L5	E_a	E_p	E_{2070}
Ageratina	*adenophora*	396	0.98	0.92	0.85	1.00	0.23	0.71	0.28	0.00	0.01	1034	1052	1382
Ailanthus	*altissima*	32	0.99	0.95	0.74	1.00	0.24	0.15	0.67	0.00	0.29	1056	1031	1417
Ambrosia	*artemisiifolia*	46	0.98	0.96	0.84	1.00	0.60	0.71	0.19	0.00	0.02	632	558	1080
Anthoxanthum	*odoratum*	338	0.96	0.84	0.90	0.96	0.76	0.08	0.35	0.00	0.01	1726	2119	2135
Carduus	*nutans*	46	0.94	0.86	0.67	1.00	0.40	0.21	0.40	0.01	0.14	1289	1646	1683
Chondrilla	*juncea*	606	0.86	0.59	0.88	0.93	0.42	0.33	0.53	0.00	0.01	1147	1356	1525
Cirsium	*vulgare*	3763	0.79	0.54	0.98	0.89	0.18	0.29	0.07	0.00	0.06	1629	1733	1829
Conium	*maculatum*	50	0.93	0.81	0.83	0.97	0.04	0.35	0.74	0.00	0.10	1244	1397	1479
Convolvulus	*arvensis*	93	0.91	0.72	0.81	0.90	0.21	0.32	0.73	0.01	0.03	1320	1240	1434
Cortaderia	*selloana*	44	1.00	0.99	0.74	1.00	0.41	0.95	0.07	0.03	0.01	303	569	837
Crataegus	*monogyna*	154	0.96	0.81	0.88	0.97	0.81	0.15	0.19	0.01	0.02	1335	1749	1726
Cytisus	*scoparius*	35	0.99	0.98	0.70	1.00	0.89	0.17	0.19	0.01	0.04	1528	1659	1629
Echium	*vulgare*	115	0.91	0.74	0.70	0.94	0.61	0.10	0.51	0.00	0.03	1297	1856	1837
Eragrostis	*curvula*	309	0.95	0.81	0.87	0.93	0.13	0.72	0.58	0.01	0.00	1333	1346	1477
Hedera	*helix*	48	0.97	0.85	0.69	0.94	0.23	0.21	0.57	0.14	0.02	1298	1948	1860
Holcus	*lanatus*	612	0.95	0.83	0.94	1.00	0.81	0.23	0.03	0.00	0.02	1594	1970	2057
Hypericum	*perforatum*	406	0.86	0.61	0.90	0.89	0.15	0.83	0.07	0.01	0.00	1390	2162	2162
Ilex	*aquifolium*	18	1.00	1.00	0.57	1.00	0.85	0.24	0.58	0.04	0.00	1036	1468	495
Leucanthemum	*vulgare*	32	0.97	0.91	0.80	1.00	0.86	0.11	0.31	0.00	0.01	1417	2116	2159
Lonicera	*japonica*	174	0.93	0.80	0.91	0.95	0.62	0.78	0.10	0.02	0.01	1332	1350	1575
Lotus	*corniculatus*	24	0.96	0.91	0.68	1.00	0.81	0.17	0.22	0.02	0.02	1350	2162	2162
Mimulus	*moschatus*	30	0.99	0.97	0.74	1.00	0.84	0.09	0.48	0.00	0.01	1591	2162	2162
Pinus	*radiata*	98	0.97	0.84	0.84	0.96	0.55	0.14	0.44	0.00	0.01	1576	1738	1861
Poa	*pratensis*	54	0.98	0.96	0.74	1.00	0.92	0.05	0.02	0.01	0.06	1336	2123	2134
Prunus	*laurocerasus*	16	0.94	0.86	-0.19	1.00	0.23	0.37	0.86	0.01	0.00	1132	2162	2162
Robinia	*pseudoacacia*	20	0.89	0.83	0.65	1.00	0.14	0.23	0.38	0.00	0.35	1110	1666	1705
Rosa	*rubiginosa*	935	0.93	0.77	0.95	0.95	0.45	0.25	0.23	0.00	0.02	1429	1562	1777
Tragopogon	*dubius*	33	0.99	0.99	0.68	1.00	1.00	0.08	0.11	0.00	0.01	1416	1711	1250
Verbascum	*thapsus*	283	0.92	0.76	0.86	0.93	0.34	0.34	0.46	0.00	0.04	1440	1603	1659

Table S5: Climatic scenarios for Switzerland at global scale.

	Abbreviation	Source	Scale	Average Temperature (°C)	Aridity
1	Current	Cgiat	30 Arc sec.	4.9 +/- 4.1	25732 +/- 16757
2	Echam5 A1 2030	Cgiat	30 Arc sec.	5.7 +/- 4.1	23369 +/- 14951
3	Hadcm3 A1 2030	Cgiat	30 Arc sec.	6.5 +/- 4.1	21935 +/- 13056
4	Ccsm3 A2 2070	Cgiat	30 Arc sec.	8.6 +/- 4.1	18702 +/- 10998
5	Hadcm3 A2 2070	Cgiat	30 Arc sec.	8.5 +/- 4.1	19718 +/- 11190
6	Echam5 A1 2070	Cgiat	30 Arc sec.	7.6 +/- 4.1	20483 +/- 12624
7	Hadcm3 A1 2070	Cgiat	30 Arc sec.	8.8 +/- 4.1	18738 +/- 10783

Table S6: climatic scenarios for New South Wales at global scale.

	Abbreviation	Source	Scale	Average Temperature (°C)	Aridity
1	Current	Cgiat	30 Arc sec.	17.0 +/- 2.6	4489 +/- 2870
2	Echam5 A1 2030	Cgiat	30 Arc sec.	18.8 +/- 2.5	3865 +/- 2960
3	Hadcm3 A1 2030	Cgiat	30 Arc sec.	18.6 +/- 2.6	3843 +/- 2650
4	Echam5 A2 2070	Cgiat	30 Arc sec.	20.6 +/- 2.6	3687 +/-
5	Hadcm3 A2 2070	Cgiat	30 Arc sec.	20.3 +/- 2.6	3586 +/- 2409
6	Echam5 A1 2070	Cgiat	30 Arc sec.	20.4 +/- 2.6	3868 +/- 2834
7	Hadcm3 A1 2070	Cgiat	30 Arc sec.	19.9 +/- 2.6	3441 +/- 2390

Table S7: Climatic scenarios for Switzerland at fine scale

	Abbreviation	Source	Scale	Average Temperature (°C)	Average annual precipitations (mm)
1	Current	Kienast & Zimmermann	100 m	5.9 +/- 3	1373 +/- 326
2	Echam5 A1 2030	WSL	100 m	7.2 +/- 2.9	1488 +/- 249
3	Hadcm3 A1 2030	WSL	100 m	7.2 +/- 2.9	1373 +/- 267
4	Ccsm3 A2 2070	WSL	100 m	8.9 +/- 2.9	1116 +/- 312
5	Echam5 A2 2070	WSL	100 m	8.3 +/- 2.8	1560 +/- 392
6	Echam5 A1 2070	WSL	100 m	9.4 +/- 2.9	1078 +/- 284
7	Hadcm3 A1 2070	WSL	100 m	10.6 +/- 3.0	1133 +/- 365

Table S8: climatic scenarios for New South Wales at fine scale.

	Abbreviation	Source	Scale	Average Temperature (°C)	Average annual precipitations (mm)
1	Current	Csiro	250 m	15.1 +/- 3.3	705 +/- 304
2	Echam5 A1 2030	Csiro	250 m	16.1 +/- 3.3	673 +/- 292
3	Hadcm3 A1 2030	Csiro	250 m	15.9 +/- 3.3	678 +/- 293
4	Echam5 A2 2070	Csiro	250 m	17.8 +/- 3.4	622 +/- 273
5	Hadcm3 A2 2070	Csiro	250 m	17.1 +/- 3.4	637 +/- 278
6	Echam5 A1 2070	Csiro	250 m	18.6 +/- 3.5	597 +/- 264
7	Hadcm3 A1 2070	Csiro	250 m	17.7 +/- 3.4	615 +/- 264

Table S9: Description of species traits and characteristics. Traits were pooled into groups for models selections and were provided by two databases Biolflor (Bf) and Flora Indicativa (Fi). Number of available species with traits information is given for Switzerland (N_{ch}) and New South Wales (N_{NSW}). The difference on each traits between Swiss and Australian species is given by a the Pue (P) of random permutation t-test for quantitative data or a Chi-squared test for qualitative data

Traits/ Characteristics	**Category**	**Levels**	**Source**	**N_{ch}**	**N_{NSW}**	**P**
Temperature	Environment	Semi-quantitative: scale from 1 (cold) to 5 (warm)	Fi	24	24	**0.040**
Continentality	Environment	Semi-quantitative: scale from 1 (oceanic) to 5 (continental)	Fi	24	24	0.074
Light	Environment	Semi-quantitative: scale from 1 (deep shade) to 5 (full light)	Fi	24	24	0.125
Soil moisture	Environment	Semi-quantitative: scale from 1 (very dry) to 5 (flooded)	Fi	24	24	0.192
Soil moisture variability	Environment	Semi-quantitative: scale from 1 (little varying) to 3 (strongly varying)	Fi	24	24	**0.026**
Soil pH	Environment	Semi-quantitative: scale from 1 (extremely acid) to 5 (alkaline)	Fi	24	24	0.402
Soil nutrients	Environment	Semi-quantitative: scale from 1 (very infertile) to 5 (over-rich)	Fi	24	24	0.137
Soil humus	Environment	Semi-quantitative: scale from 1 (little or no humus) to 5 (high humus content)	Fi	24	24	0.638
Soil aeration	Environment	Semi-quantitative: scale from 1 (bad aeration) to 5 (good aeration)	Fi	24	24	0.523
Life Form[1]	Morphology/ competitivity	Herbaceous persistent, herbaceous short-living, woody plants	Fi	24	24	0.694
Leaf duration	Morphology/ competitivity	Deciduous, partly wintergreen, wintergreen, evergreen	Fi	24	24	0.086
Life strategies	Morphology/ competitivity	Competitive, competitive ruderal, stress-tolerant competitor	Fi	24	24	**0.019**
First flowering month	Morphology/ competitivity	First flowering month in at low elevation in Switzerland	Fi	24	24	0.538
Flowering duration	Morphology/ competitivity	Flowering duration at low elevation in Switzerland	Fi	24	24	0.493
Dominance	Morphology/ competitivity	Semi-quantitative: scale from 1 (scattered) to 5 (grouped in large areas, dominating)	Fi	24	24	**<0.001**
Ploidy	Morphology/ competitivity	Diploid, more than diploid	Bf	24	24	0.353
Rosette	Morphology/ competitivity	Yes, no	Bf	22	24	0.357
Anthropochory	Dispersal/ reproduction	Yes, no	Fi	24	24	1
Meteochory	Dispersal/ reproduction	Yes, no	Fi	24	24	1
Zoochory	Dispersal/ reproduction	Yes, no	Fi	24	24	0.389
Myrmechochory	Dispersal/ reproduction	Yes, no	Fi	24	24	0.24

Boleochory	Dispersal/ reproduction	Yes, no	Fi	24	24	0.667
Vegetative dispersal[2]	Dispersal/ reproduction	Root shoots, below-ground runners, above ground runners, no vegetative dispersal, basal lateral shoots, creeping shoots or espalier	Fi	24	24	0.24
Selfing	Dispersal/ reproduction	Yes, no	Fi	24	24	0.442
Number of occurrences	Distribution	quantitative	Occ. Data	24	29	**<0.001**
Subtropical	Distribution	Yes, no	Bf	22	21	1
Meridional[3]	Distribution	Yes, no	Bf	22	21	0.067
Submeridional	Distribution	Yes, no	Bf	22	21	1
Southern Temperate[2]	Distribution	Yes, no	Bf	22	21	**0.044**
Northern temperate	Distribution	Yes, no	Bf	22	21	0.190
Boreal	Distribution	Yes, no	Bf	22	21	0.110

[1]Herbaceous chamaephytes, geophytes, biennial to several-year old hapaxanthic species and long-living hemicryptophytes were pooled in the category herbaceous persistent. Therophytes and short-living hemicryptophytes were pooled in the category herbaceous short-living and phanerophyte and nanophanerophyte were pooled in the category woody plants

[2]Tussocks and tufts were pooled with abouve ground runners

[3]Not enough variation in NSW, so not included in the NSW analyses

Table S10: coefficients of GLM's used to correlate elevation and traits. Models were applied using species observed occurrences and SDM's predictions for species found in Switzerland (CH) and New South Wales (NSW).

GLM predicting observed maximal elevation in NSW

	Estimates	Std. Error	Pr(>|t|)	
(Intercept)	765.03	168.78	0.000258	***
Moisture	228.95	62.49	0.001776	**
Anthropochory	-456.64	123.53	0.001651	**

Null deviance: 1198399 on 20 degrees of freedom
Residual deviance: 495218 on 18 degrees of freedom
AIC: 279.03

GLM predicting maximal SDM's predicted elevation in NSW under current conditions
SDM calibration : global

	Estimates	Std. Error	Pr(>|t|)	
(Intercept)	1926.7	40.14	<2e-16	***
Subtropical	-266.67	106.19	0.0212	*

Null deviance: 733782 on 20 degrees of freedom
Residual deviance: 550924 on 19 degrees of freedom
AIC: 279.27

GLM predicting SDM's predicted elevation expansion under A1b 2070 scenarios in NSW
SDM calibration : global

	Estimates	Std. Error	Pr(>|t|)	
(Intercept)	-70.1153	45.7231	0.145977	
ccs	100.1676	76.8984	0.212357	
crr	79.453	62.9794	0.226369	
crs	35.117	50.6044	0.49832	
css	-145.476	69.5705	0.053955	.
Alt. Truncation	0.5247	0.1191	0.000511	***

Null deviance: 456313 on 20 degrees of freedom
Residual deviance: 78225 on 15 degrees of freedom
AIC: 246.27

GLM predicting maximal SDM's predicted elevation in NSW
SDM calibration : local

	Estimates	**Std. Error**	**Pr(>|t|)**	
(Intercept)	1828.45	61.37	<2e-16	***
subtropical	-217.54	-2.143	0.0461	*
Zoochory	147.38	2.015	0.0591	.

Null deviance: 733782 on 20 degrees of freedom
Residual deviance: 449562 on 18 degrees of freedom
AIC: 277

GLM predicting SDM's predicted elevation expansion under A1b 2070 scenarios in NSW
SDM calibration : local

	Estimates	**Std. Error**	**Pr(>|t|)**	
(Intercept)	-353.7	119.1	0.007573	**
Deciduous	566.4	142.4	0.000739	***
Partly wintergreen	432.9	150.6	0.009397	**
Wintergreen	395.8	137.5	0.0093	**

Null deviance: 1530058 on 23 degrees of freedom
Residual deviance: 851072 on 20 degrees of freedom
AIC: 329.54

GLM predicting observed maximal elevation in CH

	Estimates	Std. Error	**Pr(>|t|)**	
(Intercept)	4388.8	639.9	1.15E-06	***
T	-716.6	150.4	0.000118	***

Null deviance: 5179970 on 21 degrees of freedom
Residual deviance: 2425337 on 20 degrees of freedom
AIC: 323.86

GLM predicting maximal SDM's predicted elevation in CH
SDM calibration : global

	Estimates	**Std. Error**	**Pr(>|t|)**	
(Intercept)	2843.3	872.8	0.00414	**
T	-477.3	146.9	0.00422	**
N	278.7	130.7	0.04627	*

Null deviance: 4200616 on 21 degrees of freedom
Residual deviance: 2078198 on 19 degrees of freedom
AIC: 322.46

GLM predicting SDM's predicted elevation expansion A1b 2070 scenarios in CH
SDM calibration : global

	Estimates	**Std. Error**	**Pr(>|t|)**	
(Intercept)	145.4913	142.2813	0.31763	
Alt. Truncation	0.3446	0.1062	0.00372	**

Null deviance: 1884921 on 23 degrees of freedom
Residual deviance: 1274991 on 22 degrees of freedom
AIC: 335.24

GLM predicting maximal SDM's predicted elevation in CH
SDM calibration : local

	Estimates	**Std. Error**	**Pr(>|t|)**	
(Intercept)	3472.4	605.9	1.97E-05	***
T	-489.1	137.9	0.00231	**
Rosettes	404.3	146.1	0.01272	*
Polyploid	-334	124.6	0.01531	*

Null deviance: 5054031 on 21 degrees of freedom
Residual deviance: 1515695 on 18 degrees of freedom
AIC: 317.52

GLM predicting SDM's predicted elevation expansion under A1b 2070 scenarios in CH
SDM calibration : local

	Estimates	**Std. Error**	**Pr(>|t|)**	
(Intercept)	287.9785	100.6603	0.00908	**
Alt. truncation	0.2682	0.05507	7.22E-05	***

Null deviance: 714876 on 23 degrees of freedom
Residual deviance: 344009 on 22 degrees of freedom

CHAPTER 2.2

Do vineyards threaten wild grapevine? A niche perspective

Blaise Petitpierre[1], Claire Arnold[1] & Antoine Guisan[1]

[1] Department of Ecology & Evolution, University of Lausanne, 1015 Lausanne, Switzerland

This paper is in preparation

My contribution to the project: I participated to the sample collections in herbaria, designed the statistical approaches, ran the niche modeling analyses and wrote the initial draft of the paper

ABSTRACT

Eurasian grapevine (*Vitis vinifera* L.) is one is one of the most important fruit crops in the world. It has been domesticated for thousand years and there is now clear morphological and habitat differences between the wild relative (*V. vinifera* ssp. *sylvestris* (Gmelin) Hegi) and the cultivars relative (*V. vinifera* ssp. *vinifera*). Since the 19th century, American *Vitis* species have been used as rootstocks to make vineyards resistant against the destructive phylloxera. These rootstocks are now naturalized and colonize similar habitats as the wild subspecies, which is now endangered. Additionally, climate change is assumed to have significant impact on grapevine distributions. In this study, we modeled the bioclimatic niche of three American rootstock species, cultivated and wild Eurasian grapevine to assess their niche overlap in Europe. We applied these models under future climate change scenarios to assess how the distribution of these species could respond to climate change. We showed a high niche overlaps between rootstocks, cultivars and the wild relative supporting an additional competition between naturalized rootstocks and wild grapevine. Additionally, each rootstock species are predicted respond differently to climate change, calling for a finer management of grape varieties for viticulture in an era of global warming. This pleads for combined efforts to protect the habitats of the wild relative in regards to rootstocks invasions and climate change, as it represents a valuable gene pool to adapt viticulture to a changing world.

INTRODUCTION

The genus *Vitis* includes about 60 species mainly distributed in the northern hemisphere. Only one species, Eurasian grapevine (*Vitis vinifera* L.) is naturally present in Europe and around the Mediterranean basin. Its domestication started somewhere in-between the Black and the Caspian Seas, at the early stage of the Neolithic period (Myles *et al.*, 2011). Since then, viticulture has rapidly diversified and spread to become one of the most widely distributed and cultivated fruit crops with important economical interests (De Mattia *et al.*, 2008, Fraga *et al.*, 2012, Arroyo García and Revilla, 2013). Despite the fact that viticulture is now present in all continents, Europe remains the world leader in term of vineyard surfaces (Castellucci, 2009).

The cultivated vine varieties (*V. vinifera* ssp. *vinifera*, hereafter abbreviated *V. vinifera*) have significantly diverged from the wild subspecies (*V. vinifera* ssp. *sylvestris* (Gmelin) Hegi, hereafter abbreviated *V. sylvestris*) and looks now very different (Levadoux, 1956, Grassi *et al.*, 2003, Aradhya *et al.*, 2003, McGovern, 2003, This *et al.*, 2006, Zecca *et al.*, 2010). Some classical authors such as Theophrastus (371BC - 287 BC) or Pliny the Elder (AD23 – AD 79) already mentioned the existence of numerous grape varieties (Bouby et al., 2013). Initially, all varieties were issued from *V. vinifera* crossings (Levadoux, 1956, Grassi *et al.*, 2003, Aradhya *et al.*, 2003, McGovern, 2003, This *et al.*, 2006) and have

remained mainly propagated vegetatively for centuries, leading to somatic mutations that also increased the number of grape varieties (Myles *et al.*, 2011).

In the middle of the 19th century, grape diseases and pests from North America were introduced in Europe (Granett *et al.*, 2001). Mainly phylloxera (*Daktulosphaira vitifoliae* Fitch) and in a lesser extent Downy mildew and powdery mildew, largely contributed to the destruction of vineyards and wild populations in most regions of Europe. As a result, lots of efforts were put since the second part of the 19th century into gathering better knowledge of wild relatives of *Vitis*. The only hope to revive vineyards in Europe depended on the ability to find vine relatives able to resist to the introduced diseases and pests, either through rootsocks or via new interspecific varieties. In America, various breeding programs were already investigating the potential use of local *Vitis* species such *as V. labrusca* L., *V. aestivalis* (Michx.) or *V. riparia* Michx. When Riley and Planchon finally discovered that *V. vinifera* could be successfully grafted onto north-american *Vitis* species (Smith, 2005), a new kind of viticulture started and propagated worldwid since the 1880s, allowing a better adaptation and regulation of the grape production. This likely also contributed to the reduction of grape varieties in the world. Indeed despite the fact that thousands of grape varieties exist, the world production is largely dominated by a few of them only, such as cabernet sauvignon, merlot or chardonnay.

The rootstock development and diversification started with Millardet, Teleki, deGrasset, Kober, Richter, and Paulsen. These names are still associated with the main rootstock varieties of the moment. The current rootstocks are mainly issue from three *Vitis* American species: *V. cinerea* (Engelm.) Engelm. ex Millard var. *helleri* (L.H. Bailey) M.O. Moore (synonymous of *V. berlandieri* Planch.), *V. riparia* Michx. and *V. rupestris* Sheele (Fig. 1). *V. riparia* has a wide distribution in North America, *V. rupestris* is mainly distributed in the southern part of the USA and *V. berlandieri* has a most restricted distribution around the Texas region. Wild american grapes are thus adapted to a wide range of climatic conditions and forest structure (Morano and Walker, 1995). The crossings were developed to be resistant to the pests and diseases, but also to be adapted to the various “terroir”.

V. sylvestris used to have a broad range of natural habitat (Arnold 2002), being distributed around the Mediterranean basin and temperate climates (Fig. 1, Heywood and Zohary, 1995, , Arroyo García and Revilla, 2013). Since the beginning of the 19th century, its distribution has dramatically regressed for two main reasons. First, it was already becoming affected by forest and river management, which intensified in the 18th century in Europe. Second, the introduction of the pests and diseases that affected the cultivated vine also contributed to the extermination of the wild parent in habitats where the diseases could spread. These factors drove its status close to extinction, so as many wild varieties of other cultivated plants (Heywood and Zohary, 1995), in particular at the western side of its distribution (IUCN 1997, Arnold *et al.*, 1998, Arnold, 2002, Arnold *et al.*, 2005, Di Vecchi-Staraz *et al.*, 2009, Arroyo García and Revilla, 2013, Red list in CH, Red list in France and Spain). Nowadays, wild grapevine is mainly restricted to floodplain habitats.

There have been evidences of escapes from cultivated vineyards towards natural habitats similar to the ones occupied by the European wild grapevine (Laguna, 2004, Arrigo and Arnold, 2007). These rootstocks are vectors of diseases and pests. Due to their specific selection, they also proved to be serious competitors in floodplain habitat (Arrigo and Arnold, 2007). Rootstocks may also represent a genetic invader (especially in the inbred and declining populations in western Europe), as they have a remarkable ability to hybridize with sister species in their native environment (Levadoux, 1956, Reynaud, 1971, Zecca *et al.*, 2012). Although evidence of geneflow between rootstocks and wild wine have been limited so far (Arrigo and Arnold, 2007, Zecca *et al.*, 2010), interbreeding between cultivated and wild wine exists (Di Vecchi-Staraz *et al.*, 2009, De Andrés *et al.*, 2012) and may affect future evolution of the wild variety.

While the genetic relatedness between cultivars and within the *Vitis* genus has been intensively studied (e.g. Bowers *et al.*, 1999, Grassi *et al.*, 2003, Aradhya *et al.*, 2003, Arroyo-García *et al.*, 2006, This *et al.*, 2006, Zecca *et al.*, 2010, De Andrés *et al.*, 2012, Zecca *et al.*, 2012, Poczai *et al.*, 2012), the relationships between their macroclimatic niches has attracted much less attention, despite its importance for assessing future areas of grapewine production in a context of rapide climatic changes, with potentially huge economic and cultural implications (Kenny and Harrison, 1992, Nemani *et al.*, 2001, Jones *et al.*, 2005, Meier *et al.*, 2007, White *et al.*, 2006, Hannah *et al.*, 2013). In particular, assessing whether a distribution overlap can lead to possible genetic introgression preliminary requires assessing niche overlap. Moreover, rapid climate change represents an important threat for the biodiversity and the persistence of species, especially those limited in their dispersal ability (e.g. Dullinger *et al.*, 2012). Ecological niche modeling (ENM) is a simple and efficient way to model species ecological niches by relating species distribution to environmental variables (Guisan and Zimmermann, 2000, Peterson and Vieglais, 2001, Elith and Leathwick, 2009). If environmental variables are spatially explicit, it is then possible to project the fitted ENMs in geographical space, resulting in the spatial prediction of the species‘potential distribution. With ENMs, it is then possible to assess how potential distributions can be affected by climate or landuse change scenarios, provided that the latter are available for the variables used in the ENMs, thus providing a potentially very useful information for conservation managers (Guisan and Thuiller, 2005, Guisan *et al.*, 2013). Additionally, it is possible to compare, quantify and statistically test the ecological niche similarity between different species (e.g. in an evolutionary perspective), or for a same species between its native and non-native ranges (Warren *et al.*, 2008, Broennimann *et al.*, 2012). Regarding the wild grapevine, modeling its niche and the ones of its cultivated varieties could thus provide a better understanding of their current and future distributions, and deliver insightful information to increase the success of conservation plans for the wild and endangered variety (Arnold *et al.*, 2005).

Here, we aim at quantifying the degree of niche overlap between wild wine, cultivated vineyards and American *Vitis* species used as rootstock, and assess the risk of the wild type being threatened by viticulture expanding on its habitats or by rootstock species

naturalizing and invading its suitable habitats. To date, such comparisons have never been done comprehensively, especially at a scale large-enough to capture full species' niches, including rootstocks. For this purpose, we modeled the niche of three American species used as the main rootstocks in Europe (*V. Berlandieri*, *V. riparia* and *V. rupestris*), of the European vineyards and of the wild wine. We quantified the niche of each species, assessed niche overlap between all species pairs and projected the potential distribution of each species in Europe under current and future climatic conditions. This allowed assessing how viticulture may occupy suitable habitats for wild wine now and under a future warmer climate.

METHODS

Species occurrences and environmental data

Species distribution data were gathered using several databases. For all species, the GBIF database (data.gbif.org) was screened in August 2013. Only georeferenced data with accuracy below 20 km were kept for further analyses. Additionally, we included information of every *V. sylvestris* specimen obtained from the herbaria of Geneva, Lausanne and Zürich when georeferences were available. We also included distribution data from the AgroAtlas database (Afonin *et al.*, 2008) for *V. sylvestris* and *V. vinifera* in Eastern Europe thus compiling the most exhaustive distribution database for *V. sylvestris* to date. For distribution of vineyard in Western Europe, we used the CORINE land cover (The European Topic Centre on Land Use and Spatial Information available on http://www.eea.europa.eu/data-and-maps/data/corine-land-cover-2006-raster#tab-gis-data). South-eastern Europe was complemented with the maps from Johnson (2002). Information about 33 European occurrences of escaped rootstocks came from personal observation and personal communications (Occete) and Arrigo and Arnold, 2007, Fig. 1).

Environmental data were gathered from the CliMond database (Kriticos *et al.*, 2011). It consists of an environmental dataset composed by 35 variables at a resolution of ca. 16 km, grouped into 4 categories: temperature, precipitation, moisture and solar radiations. We did not include solar radiations as this information was already used in the calculation of soil moisture variables, a more proximal factor to predict plant species distribution at this scale (Table 1). Two sets of variables were used in the species distribution models (ENMs, see below). The first set (V_{env}) included 8 variables commonly used to predict naturalized species distribution (Table 1, e.g. Thuiller *et al.*, 2005, Broennimann *et al.*, 2007) and was shown to maximize ENM transferability across geographic regions (Petitpierre *et al.,* in prep.). As only few occurrences were available for *V. berlandieri* and *V. rupestris*, we summarized the environmental variation available in Europe (EU) and North America (NA) through a principal component analysis (PCA, see below) and included only the two first components as predictors (V_{PCA}). This approach was shown to be a robust way to predict species exotic ranges in the Holarctic kingdom (Petitpierre *et al.*, in prep.) with the advantage of not overparametrizing ENMs calibrated with few occurrences. Although ENM

were calibrated with these two distinct sets of predictors , we only present predictions obtained with V_{env} for *V. riparia*, *V. sylvestris* and *V. vinifera* and V_{pca} for *V. berlandieri* and *V. rupestris* (see Fig. S1 for predictions with the reciprocal environmental dataset). Two global climate circulation models (GCM, CSIRO-MK3·0 and MIROC-H) for two climate change scenarios (based on the 4th assessment of the Intergovernmental Panel on Climate Change; A1B and A2 IPCC, 2007) were available for the same climatic variables in the CliMond database.

Table 1: Variables used to calibrate the principal component analysis. Variables in bold are the one used in the species distribution models (V_{env})

Variable Number	Variable
Bio01	**Annual mean temperature (°C)**
Bio02	Mean diurnal temperature range (mean(period max-min)) (°C)
Bio03	Isothermality (Bio02 ÷ Bio07)
Bio04	**Temperature seasonality (C of V)**
Bio05	Max temperature of warmest week (°C)
Bio06	Min temperature of coldest week (°C)
Bio07	Temperature annual range (Bio05-Bio06) (°C)
Bio08	Mean temperature of wettest quarter (°C)
Bio09	Mean temperature of driest quarter (°C)
Bio10	**Mean temperature of warmest quarter (°C)**
Bio11	**Mean temperature of coldest quarter (°C)**
Bio12	Annual precipitation (mm)
Bio13	Precipitation of wettest week (mm)
Bio14	Precipitation of driest week (mm)
Bio15	**Precipitation seasonality (C of V)**
Bio16	**Precipitation of wettest quarter (mm)**
Bio17	Precipitation of driest quarter (mm)
Bio18	Precipitation of warmest quarter (mm)
Bio19	Precipitation of coldest quarter (mm)
Bio28	**Annual mean moisture index**
Bio29	Highest weekly moisture index
Bio30	Lowest weekly moisture index
Bio31	**Moisture index seasonality (C of V)**
Bio32	Mean moisture index of wettest quarter
Bio33	Mean moisture index of driest quarter
Bio34	Mean moisture index of warmest quarter
Bio35	Mean moisture index of coldest quarter

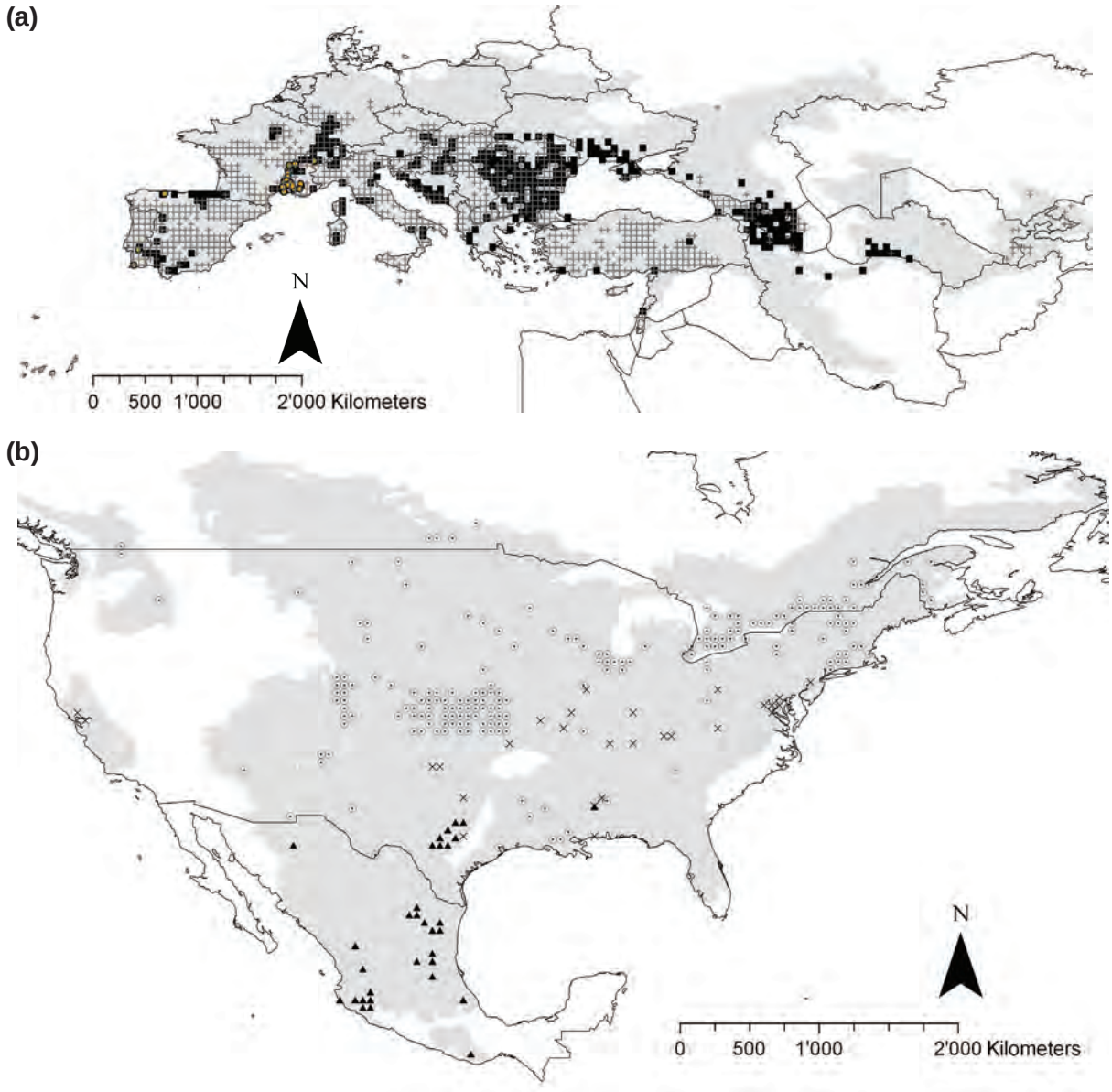


Figure 1: Study areas and occurrences used to calibrate models. In Europe (EU, a), *V. sylvestris* and *V. vinifera* occurrences are represented by black square and cross respectively. Additionally, occurrences of naturalized rootstocks are symbolized with orange dots. In North America (NA, b), are *V. berlandieri*, *V. riparia* and *V. rupestris* occurrences are reprensented by triangle, dotted circle and cross repectively. Grey area show the backgrounds (also called “study area”) used to calibrate ENM in each study area.

The delimitation of the extent used in ENMs may have severe impact on their geographic predictions or niche overlap assessment and should correspond to a biogeographically relevant area in regard to the studied species (Barve *et al.*, 2011, Rödder and Engler, 2011). The two study areas were then defined as the union of the ecoregions covered by the three rootstocks species in NA and respectively the ecoregions covered by wild wine and vineyards in EU (Fig. 1). Ecoregions are defined as relatively large units of land containing a distinct assemblage of natural communities sharing a large majority of species, dynamics, and environmental conditions (Olson *et al.*, 2001).

Principal component analysis and test of niche conservatism

A principal component analysis (PCA) was calibrated on the 27 climate variables covering the pooled extents in NA and EU. We used the two first components as environmental predictors in ENM (see below) as a way to depict the environmental space to assess niche overlap. This approach (referred as PCA_{env} in Broennimann *et al.*, 2012) was shown to be the most accurate to assess species niche overlap in Holarctic climate (Broennimann *et al.*, 2012). Species occurrences and background extents were then projected into this environmental space, which was then gridded (resolution R = 100) to calculate species occupancies smoothed by a simple kernel function (Broennimann *et al.*, 2012). Occupancies reflect species densities distribution in the environmental space, corrected by the general availability of the different environments. We used it to compare the niche overlap between the 5 taxa. Additionally, we pooled occurrences of the three American rootstocks and considered them as a 6^{th} taxa. Overlap was assessed as the Schoener's D between occupancies of two taxa. The overlap D can vary between 0 (no overlap) and 1 (complete overlap). Additionally, we tested if the overlap was significantly higher than random through a one tailed similarity test based on Schoener's D (Warren *et al.*, 2008, Broennimann *et al.*, 2012).

ENM geographical predictions

We used an ENSEMBLE modeling approach combining several statistical techniques (Araujo and New, 2007). Output of three techniques commonly used in ENMs were averaged: generalized linear models (GLM, McCullagh and Nelder, 1989), gradient boosting machines (GBM, Friedman, 2001) and maximum entropy regression (MaxEnt, Phillips *et al.*, 2006). 10'000 pseudo-absences were sampled in the study area and weighted such that presences and pseudo-absences have the same prevalence in ENM. To lower the potential impact of sample bias and spatial autocorrelation (Dormann, 2007), species occurrences were disaggregated keeping a minimal distance of 0.5° (about 50 km.). Disaggregation is analogous to occurrences thinning (Verbruggen *et al.*, 2013). The whole data set was split into a 70% partition for calibrating ENMs and a 30% partition for pseudo-independent evaluation. ENM were evaluated using 4 indices: 1) Area Under the Curve of a Receiver Operating Characteristics (AUC, Zweig and Campbell, 1993), True Skill Statistics (TSS, also known as corrected Kappa, Allouche *et al.*, 2006) and two "presence-only" evaluators, the Boyce index (B, Hirzel *et al.*, 2006) and the rate of correctly predicted presences with binarized predictions, also called sensitivity (Se). For rarer species limited in their dispersal, such as *V. sylvestris,* presence-only evaluators can be particularly relevant as they do not include the rate of false positive in the evaluation. AUC varies between 0 (counter prediction) 0.5 (random model) and 1 (perfect fit) whereas TSS and B are scaled between -1 and 1. Continuous predictions were binarized using the threshold providing the best TSS to asses Se. Variables importance was assessed through randomization (see the documentation file of the BIOMOD package).

While ENMs were calibrated on the native extent of each taxa, they were then projected onto the whole Europe (EU). Additionally, ENMs were projected in 2050 and 2100, using two climate scenarios - A1B and A2 – derived from the two GCMs. We built a consensual map of the American rootstock species in EU by selecting the highest suitability among the three species for each pixel in EU. Potential geographical overlap between wild grapevine, vineyards and American rootstocks under current and future conditions was assessed using a Pearson correlation between the predictions of each taxa. Species distribution modeling was done using the R software (R, 2005) using the package BIOMOD for GLM and GBM (Thuiller *et al.*, 2009) and dismo for MaxEnt (Hijmans, 2010). For each spatial and temporal projections, we avoided predicting to non-analog climate because ENM extrapolation towards climates non-existing in the calibration range are highly hazardous (Fitzpatrick and Hargrove, 2009). We defined non-analog climate as those conditions where the value of one of the variables used to fit the ENMs was not comprised within the range covered in the calibration range. The whole procedure (calibration, evaluation, projections) was replicated 10 times per species and averaged.

RESULTS

Principal component analysis and niche conservatism

The two first axes of the PCA explain 43 % and 26 % of the total variation within the pooled NA and EU calibration extents (Fig. 2). The first axis corresponds to the aridity gradient whereas the second axis is rather related to temperature variation and cold temperatures (Fig. 2).

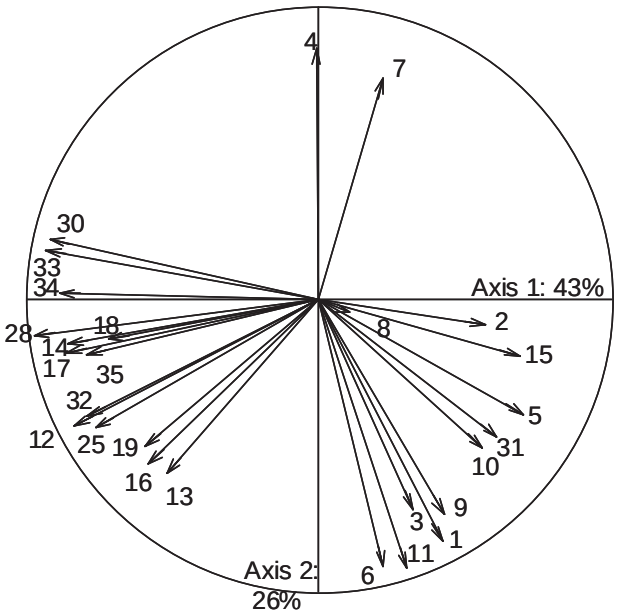


Figure 2: PCA correlation circle showing how the two first components are correlated to each variable and how much they explain total climate variation. Numbers represent the climatic variables and are defined in Table 1 (with ‘bio’ prefixes before each number).

The niches of *V. sylvestris* and *V. vinifera* show a significant and high overlap (Table 2). Both are centered in similar climate conditions and share the same niche limits (Fig. 3). The niche of the three American rootstock species pooled together also shows a significant (or marginally significant) high overlap with both European *Vitis* (Table 2), with niche more expanded along the two extremities of axis 2, i.e. towards colder and hotter temperature minima and maxima respectively, and towards higher temperature and moisture seasonality. Note that the expansion towards hotter climates and more moisture seasonality occurs in climate that is available only in NA, not in EU.

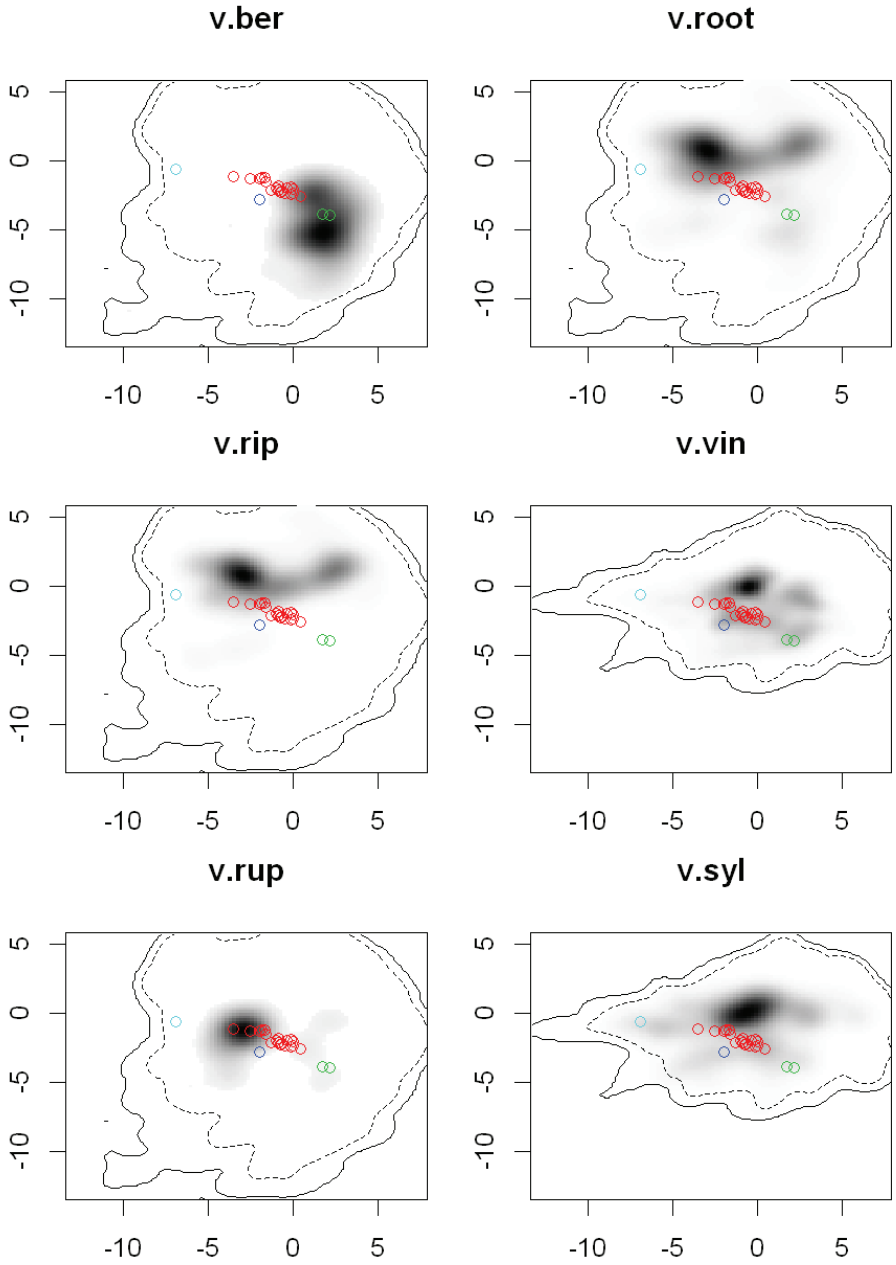


Figure 3: Environmental space depicted by the two first axes of a PCA calibrated on climate variables distributed in NA and EU study areas. Black lines show the distribution of the available climate in both extents (with dashed line representing the limit of the 75th quantile of the overall climatic density). Species distribution in the environmental space is represented by the cloud densities. Naturalized rootstocks populations from Switzerland (turquoise), France (red), Spain (blue), Portugal (green) were also projected into the environmental space.

Table 2: Niche overlap (Shoener's D) between each species and significance of the niche similarity test. When two niches are compared, the niche similarity test randomizes only one of the two niches at the same time. In the table, species displayed as rows are the one randomized.

	V. ber	*V. rip*	*V. rup*	*V. root*	*V. syl*	*V. vin*
V. ber		0.040	0.186	0.285***	0.190*	0.344**
V. rip	0.040*		0.301	0.700***	0.401	0.294
V. rup	0.186***	0.301		0.363***	0.329*	0.398**
V. root	0.285***	0.700**	0.363		0.535**	0.450'
V. syl	0.190**	0.401	0.329'	0.535**		0.501**
V. vin	0.344***	0.294	0.398*	0.450***	0.501**	

When the niches of the three American species are taken individually, they occupy distinct subniches. *V. blerlandierii* is the one occurring in dryer and hotter climates whereas *V. rupestris* and *V. riparia* occupy a wider range of moisture but confined to colder and more seasonal climates. The niche of *V. berlandieri* is confined to more specific climate, which is translated by lower overlap with the other rootstock species. On the other hand, the high overlap between the niche of *V. riparia* and the pooled rootstocks niche show that *V. riparia* is the main component of the rootstocks niche.

Predictions of ENM

ENM performances vary between poor and excellent, depending on the evaluation index used and the species considered (Table 3). *V. rupestris* shows a bad B value with both predictor datasets, whereas the other evaluators are good to excellent for this species. This suggests that the correlation between predictions and observed occurrences is low but enough to discriminate presences from absences. *V. berlandieri* also show poorer B with the V_{env} dataset (0.36 ±0.199). Nevertheless, the values of the other evaluators are rather high, in particular with the V_{PC} dataset. *V. sylvestris* has poor presence/absence but good presences-only evaluators, translating a high rate of false positives.

Table 3: Evaluation of the ENSEMBLE species distribution models for each species calibrated with 8 environmental predictors (env) or with the two first components of the PCA (pca). Additionally, the number of occurrences used to calibrate the model (Nocc), after disaggregation, is provided.

Sp	**Nocc**	**AUC_{env}**	**TSS_{env}**	**$Boyce_{env}$**	**Sen_{env}**	**AUC_{pca}**	**TSS_{pca}**	**$Boyce_{pca}$**	**Sen_{pca}**
V. ber	17	0.94+/- 0.019	0.833+/- 0.064	0.36+/- 0.199	1+/-0	0.938+/- 0.072	0.841+/- 0.123	0.442+/- 0.230	0.95+/- 0.112
V. rip	98	0.795+/- 0.03	0.495+/- 0.058	0.892+/- 0.055	0.752+/- 0.064	0.77+/- 0.025	0.484+/- 0.039	0.836+/- 0.057	0.733+/- 0.062
V. rup	17	0.828+/- 0.1	0.654+/- 0.174	-0.005+/- 0.268	0.94+/- 0.077	0.85+/- 0.045	0.655+/- 0.069	0.368+/- 0.398	0.955+/- 0.073
V. syl	140	0.771+/- 0.021	0.453+/- 0.035	0.909+/- 0.062	0.858+/- 0.08	0.694+/- 0.026	0.320+/- 0.051	0.588+/- 0.210	0.776+/- 0.116
V. vin	90	0.768+/- 0.018	0.465+/- 0.034	0.937+/- 0.045	0.824+/- 0.044	0.78+/- 0.0267	0.442+/- 0.053	0.841+/- 0.151	0.707+/- 0.062

Temperature is more important for American than for European species, which niche is rather shaped by temperature seasonality, as shown for *V. berlandieri* (Fig. 4). The moisture index seasonality is more important for *V. sylvestris* than for *V. vinifera* (Fig. 4). When ENMs are built with the first two PCA components, the second axis has a more important contribution for all taxa (Fig. 4).

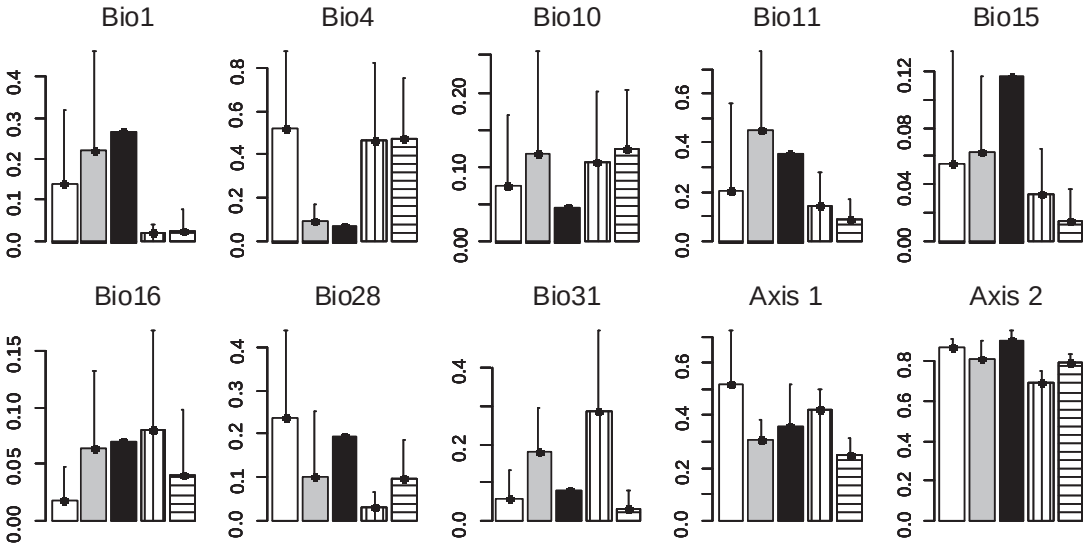


Figure 4: Variable importance in ENM for *Vitis berlandieri* (white), *V. riparia* (grey), *V. rupestris* (black), *V. sylvestris* (vertical lines) and *V. vinifera* (horizontal lines). The climatic variables are abbreviated as in Table 1.

The niche overlap illustrated through the PCA environmental space is also translated at the geographical level. When predictions for American rootstocks are pooled together, they show a wider suitable area in EU than cultivated and wild grapevine (Fig. 5). The naturalized populations of escaped rootstocks are predicted with high suitability values by the pooled predictions (Fig. 5). The correlations between predictions of American rootstocks, wild and cultivated grapevine in EU vary between 0.6 (*V. sylvestris* and American rootstocks) and 0.9 (*V. sylvestris* and *V. vinifera*) under current conditions (Fig. 6). These correlations remain stable following the different climatic scenarios, except with the MIRO's A2 2050 and A2 2100 scenarios, where the correlation between American rootstocks and wild and cultivated grapevine drops below 0.5 (Fig. 6). Potential distribution of *V. sylvestris* in EU is also wider than its current observed distribution, potentially translating an important range unfilling. This unfilling occurs within the species' observed distribution and also in Scandinavia and England, areas that the species never reached. Despite the high overlap between *V. vinifera* and *V. sylvestris*, some areas are (and remain also under climate change scenarios) more favourable to *V. sylvestris* among its current distribution. These sites are distributed between central and eastern EU in Austria, Romania and Caucasus (Fig. 7).

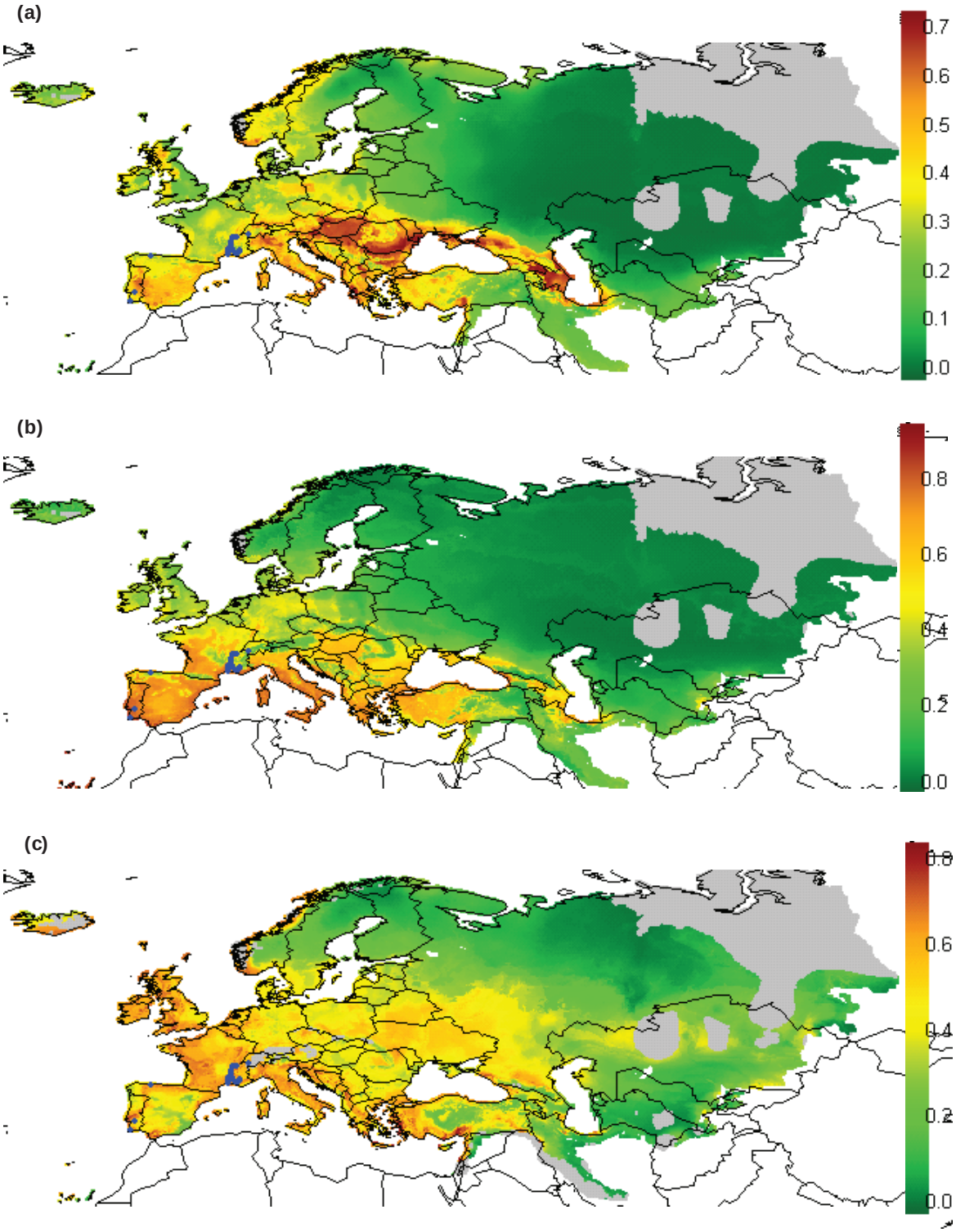


Figure 5: potential distribution of *V. sylvestris* (a*), V. vinifera* (b) and american rootstocks (c) in EU under current conditions. Occurrences of escaped and naturalized rootstocks are shown with blue dots. The color scheme indicates climatic suitability from low (green) to high (darkred).

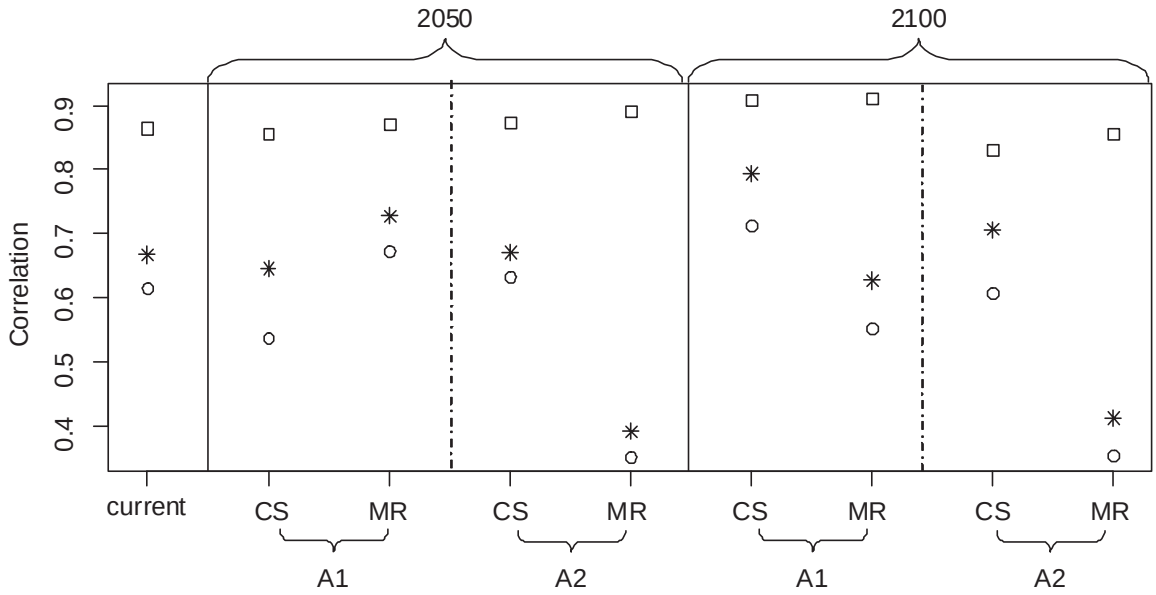


Figure 6: correlation between spatial predictions of *Vitis vinifera* and *V. sylvestris* (squares), *V. sylvestris* and american rootstocks (circles) and *V. vinifera* and american rootstocks (stars) under different climate scenarios in Europe (A1 and A2) for two GCMs (CSIRO and MIRO) at two time steps (2050 and 2100).

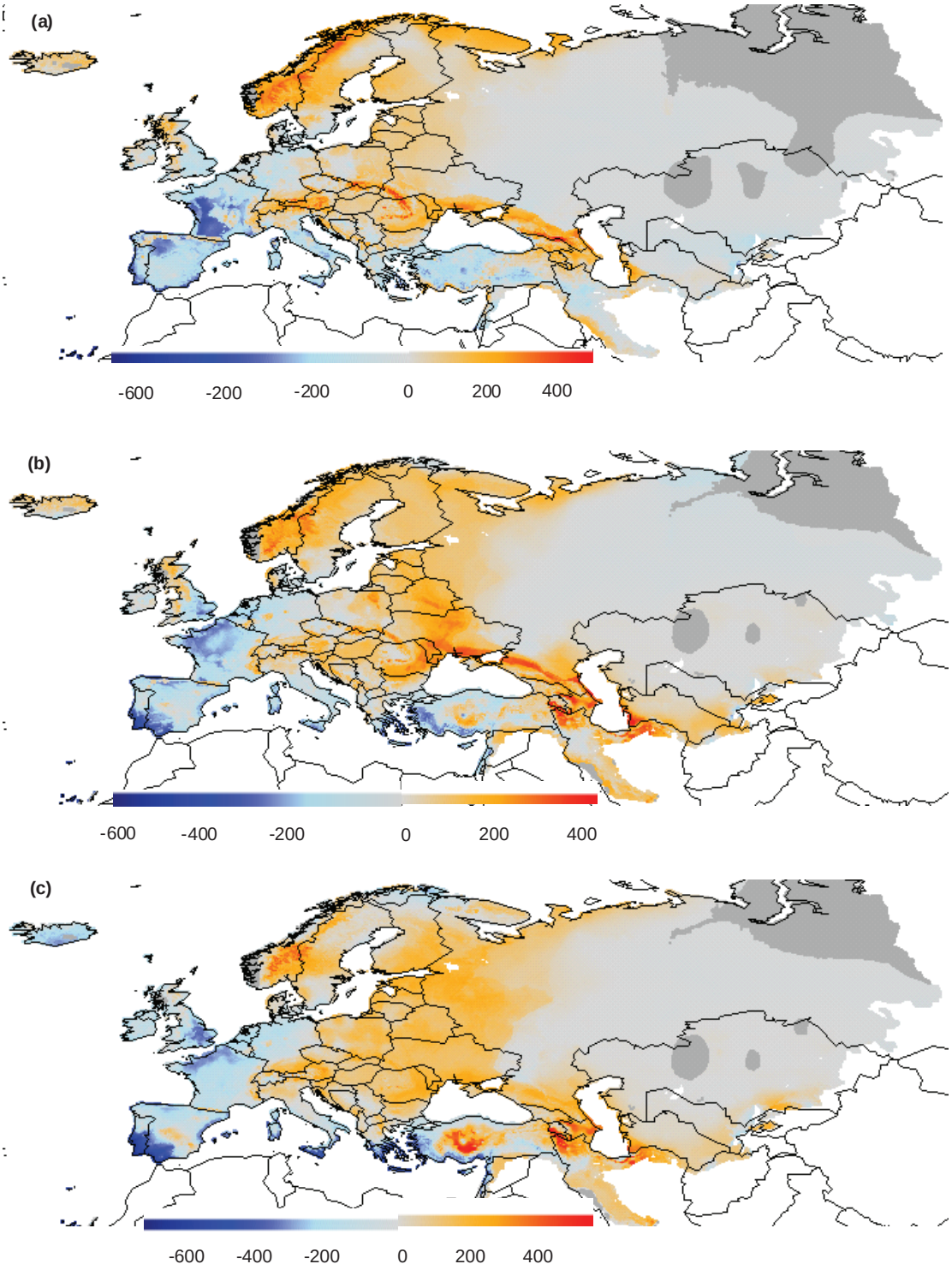


Figure 7: differences between ENM predictions for V. sylvestris and V. vinifera under current (a), A1B 2050 (b) and A1B 2100 (c) climate conditions. High values (red) show areas more favorable to V. sylvestris than to viticulture, and low values (blue) show the reverse (i.e. more favourable to viticulture). Predictions for both GCMs were averaged for each reference time.

DISCUSSION

This study depicts the climate requirements at the biogeographical level of grapevine, in contrast to the cultivated grape varieties, the wild ancestor and the rootstock sister species. We showed in this regard that high bioclimatic niche overlaps exist between American rootstocks, cultivated and wild grapevine, which also translates into an important potential geographical overlap in Europe. This highlights locations in Europe where viticulture and naturalized rootstocks could compete, now and in the future, for the same climatic conditions than the threatened wild type *V. sylvestris*.

Niche overlap

Both wild and cultivated European grapevines share a similar climatic niche as shown by Fig. 3 and Table 2. This is not surprising as the divergence of these two taxa started only recently with human history. Their niche difference occurs mostly at the habitat level where alluvial and colluvial forests are more suitable for *V. sylvestris* (Arnold *et al.*, 1998). Their main bioclimatic difference occurs on moisture seasonality which appears to be lower for *sylvestris* (Fig. 3 and 4).

American rootstock species have lower niche overlap among each other than they have with both wild and cultivated European grapevine. Thus, each of the American species occupies different parts of the environmental space. *V. berlandieri* is located in climate characterized by a high seasonality, with hotter and dryer extremes. These macroclimatic conclusions corresponds to species' ecology at the habitat scale, where *V. berlandieri* grows in the driest conditions compared to the two other species (Morano and Walker, 1995). Projected in Europe, it corresponds to the more Mediterranean part of the grapevine distribution (Fig. S1). As shown by the highest overlap with the overall rootstock niche, *V. riparia* grows in the widest environmental bioclimatic range, thus covering most of the overall niche of rootstocks, except the part occupied by *V. berlandieri*. This wide coarse bioclimatic niche may be driven at finer scale by preference for sandy soils (Morano and Walker, 1995). Finally, *V. rupestris* occurs in median conditions, overlapping with both the niche of *V. berlandieri* and *V. riparia* although it occupies the wettest conditions at the habitat level (Morano and Walker, 1995). Together, they present a wider niche, highly and significantly overlapping with the ones of the Eurasian wild and cultivated grapevines (Fig. 3, Table 2) supporting a high climatic suitability for the naturalized rootstocks in Europe. This climatic compatibility with the niche of European grapevine adds to a habitat compatibility with wild grapevine, where naturalized rootstocks successfully outcompete *V. sylvestris* (Arrigo and Arnold, 2007), whereas *V. vinifera* do not.

Potential distributions under current and future climates

The niche overlap is also translated at the geographical level (Fig. 5 and 7). Most of the actual (Fig. 1) and potential distribution of *V. sylvestris* is overlapped by *V. vinifera* and the potential distribution of rootstocks is even wider in Europe (Fig. 5 and 7). The bioclimatic conditions of the escaped rootstocks match distinctively with the native niche of rootstocks. Swiss population is closer to the niche of *V. riparia*, French and Spanish populations correspond to *V. rupestris* and *riparia* whereas Portuguese populations fall within the niche of *V. berlandieri* (Fig. 3). This illustrates how possible different combinations of rootstocks can be adapted to different climatic conditions and grape varieties.

These evidences for an additional competitive pressure on an endangered wild subspecies plead for conservation measures (Arnold *et al.*, 2005). Mapping areas where *V. sylvestris* could grow with less pressure from viticulture would allow a more efficient preservation. Our results highlight two areas in central Europe and Caucasus where ENM's suitabilities are more favorable to the endangered wild ecotype (Fig. 7). More than bioclimatic refuges, these two regions represent two distinct gene pools within *V. sylvestris* (Myles *et al.*, 2011) and thus can be prioritize for conservation efforts. However, climatic suitability could change in the future, possibly affecting both the distribution of viticulture (Hannah *et al.*, 2013) and conservation planning (Araújo *et al.*, 2004). Interestingly, these two spots keep a higher suitability for *V. sylvestris* with the projections under climate scenarios (Fig. 7 and S2).

When applied with climatic scenarios, ENMs' predictions show important changes in potential distributions, affecting rootstocks, cultivated and wild grapevines (Fig. S1). They show an overall increased suitability for wild and cultivated grapevine in Europe, with the release of novel suitable areas in north-eastern Europe (Fig. S1). Our predictions for cultivated grapevine in 2050 corroborate with the one of a recent study about future potential of viticulture (Hannah *et al.*, 2013). The high geographical correlation between *V. sylvestris* and *V. vinifera* is not predicted to change dramatically in the future (Fig. 6). However, two climate scenarios (A2_2050_MR and A2_2100_MR) show a discrepancy between potential rootstocks distributions and European wild and cultivated grapevines (Fig. 6). This discrepancy also occurs among the three American species. While the suitability is predicted to globally increase for *V. riparia* and *V. rupestris* in Europe, potential distribution of *V. berlandieri* is predicted to shrink (Fig. S1). Such different responses suggest that novel grape varieties may be necessary to adapt viticulture, to climate change, especially to replace varieties currently adapted to southern Europe. For such purpose, conserving a high genetic diversity within wild relative is crucial (Kell *et al.*, 2012). With a higher diversity compared to cultivars, *V. sylvestris* can be protected as a phylogenetic resource (Myles *et al.*, 2011, Arroyo García and Revilla, 2013). Preserving populations of wild grapevine appears to be easier if climate suitability increases as predicted for *V. sylvestris* in Europe (Fig. S1). However, due to its low dispersal ability, especially in the fragmented and senescent population on the western distribution (Di

Vecchi-Staraz *et al.*, 2009), this should not particularly benefits the species pleading again for active conservation programs.

CONCLUSION

For the first time, the bioclimatic niche of Eurasian grapevine was investigated in regards to the wild relative, the cultivars and the American species used as rootstocks, underlying the complexity behind the vineyards. Viticulture is facing distinct components of global change. First, due to an important niche overlap, the endangered wild relative is facing the competition of naturalized rootstocks able to grow in similar habitats. Second, climate warming is predicted to strongly modify the distribution of viticulture but affecting differently rootstock and Eurasian species. Such disruption will probably require novel variety, for which a highly diverse gene pool is desirable. This supports and provides tools for active and conservation measures to protect *V. sylvestris*. Modeling the niche at a finer habitat scale, where other relevant factors such as soil quality or light availability could be included is a promising way to assess niches relationships between these taxonomic units and understand both their current und future distributions.

REFERENCES

Afonin, A.N., Greene, S.L., Dzyubenko, N.I. & Frolov, A.N. (2008) http://www.agroatlas.ru

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

Aradhya, M.K., Dangl, G.S., Prins, B.H., Boursiquot, J.-M., Walker, M.A., Meredith, C.P. & Simon, C.J. (2003) Genetic structure and differentiation in cultivated grape, *Vitis vinifera* L. *Genetical research,* **81,** 179-192.

Araújo, M.B., Cabeza, M., Thuiller, W., Hannah, L. & Williams, P.H. (2004) Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology,* **10,** 1618-1626.

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution,* **22,** 42-47.

Arnold, C. (2002) *Ecologie de la vigne sauvage en Europe:(Vitis vinifera ssp. silvestris).* vdf Hochschulverlag AG.

Arnold, C., Gillet, F. & Gobat, M. (1998) Situation de la vigne sauvage *Vitis vinifera* ssp. *sylvestris* en Europe. *Vitis,* **37,** 159-170.

Arnold, C., Schnitzler, A., Douard, A., Peter, R. & Gillet, F.o. (2005) Is there a future for wild grapevine (Vitis vinifera subsp. silvestris) in the Rhine Valley? *Biodiversity & Conservation,* **14,** 1507-1523.

Arrigo, N. & Arnold, C. (2007) Naturalised *Vitis* rootstocks in Europe and consequences to native wild grapevine. *Plos one,* **2,** e521.

Arroyo-García, R., Ruiz-García, L., Bolling, L., Ocete, R., López, M.A., Arnold, C., Ergul, A., Söylemezogu, G., Uzun, H.I., Cabello, F., Ibáñez, J., Aradhya, M.K., Atanassov, A., Atanassov, I., Balint, S., Cenis, J.L., Costantini, L., Gorislavets, S., Grando, M.S., Klein, B.Y., McGovern, P.E., Merdinoglu, D., Pejic, I., Pelsy, F., Primikirios, N., Risovannaya, V., Roubelakis-Angelakis, K.A., Snoussi, H., Sotiri, P., Tamhankar, S., This, P., Troshin, L., Malpica, J.M., Lefort, F. & Martinez-Zapater, J.M. (2006) Multiple origins of cultivated grapevine (*Vitis vinifera* L. ssp. *sativa*) based on chloroplast DNA polymorphisms. *Molecular ecology,* **15,** 3707-3714.

Arroyo García, R.A. & Revilla, E. (2013) The Current Status of Wild Grapevine Populations (Vitis vinifera ssp sylvestris) in the Mediterranean Basin. *The Mediterranean Genetic Code - Grapevine and Olive* (ed B. Sladonja).

Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberon, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling,* **222,** 1810-1819.

Bouby, L., Figueiral, I., Bouchette, A., Rovira, N., Ivorra, S., Lacombe, T., Pastor, T., Picq, S., Marvinval, P., Terral, J. F. (2013). Bioarchaeological Insights into the Process of Domestication of Grapevine (*Vitis vinifera* L.) during Roman Times in Southern France. *PloS one*, *8*(5), e63195

Bowers, J., Boursiquot, J.-M., This, P., Chu, K., Johansson, H. & Meredith, C. (1999) Historical genetics: the parentage of Chardonnay, Gamay, and other wine grapes of northeastern France. *Science,* **285,** 1562-1565.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters,* **10,** 701-709.

Castellucci, F. (2009) World statistics. *7th General Assembly of the OIV. June 28th July 3rd. Zagreb, Croatia. International Organisation of Vine and Wine (OIVV),* **41**.

De Andrés, M., Benito, A., PéREZ-RIVERA, G., Ocete, R., Lopez, M., Gaforio, L., Muñoz, G., Cabello, F., MARTíNEZ ZAPATER, J.M. & ARROYO-GARCÍA, R. (2012) Genetic diversity of wild grapevine populations in Spain and their genetic relationships with cultivated grapevines. *Molecular ecology,* **21,** 800-816.

De Mattia, F., Imazio, S., Grassi, F., Baneh, H.D., Scienza, A. & Labra, M. (2008) Study of genetic relationships between wild and domesticated grapevine distributed from middle east regions to European countries. *Rendiconti Lincei,* **19,** 223-240.

Di Vecchi-Staraz, M., Laucou, V.r., Bruno, G.r., Lacombe, T., Gerber, S., Bourse, T., Boselli, M. & This, P. (2009) Low Level of Pollen-Mediated Gene Flow from Cultivated to Wild Grapevine: Consequences for the Evolution of the Endangered Subspecies *Vitis vinifera* L. subsp. *silvestris*. *Journal of Heredity,* **100,** 66-75.

Dormann, C.F. (2007) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography,* **16,** 129-138.

Dullinger, S., Gattringer, A., Thuiller, W., Moser, D., Zimmermann, N.E., Guisan, A., Willner, W., Plutzar, C., Leitner, M., Mang, T., Caccianiga, M., Dirnbock, T., Ertl, S., Fischer, A., Lenoir, J., Svenning, J.-C., Psomas, A., Schmatz, D.R., Silc, U., Vittoz, P. & Hulber, K. (2012) Extinction debt of high-mountain plants under twenty-first-century climate change. *Nature Clim. Change,* **2,** 619-622.

Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology Evolution and Systematics,* **40,** 677-697.

Fitzpatrick, M.C. & Hargrove, W.W. (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation,* **18,** 2255-2261.

Fraga, H., Malheiro, A., Moutinho-Pereira, J. & Santos, J. (2012) An overview of climate change impacts on European viticulture. *Food and Energy Security,* **1,** 94-110.

Friedman, J.H. (2001) Greedy function approximation: a gradient boosting machine. *The Annals of Statistics,* **29,** 1189-1232.

Granett, J., Walker, M.A., Kocsis, L. & Omer, A.D. (2001) Biology and managment of grape phylloxera. *Annual Review of Entomology,* **46,** 387-412.

Grassi, F., Labra, M., Imazio, S., Spada, A., Sgorbati, S., Scienza, A. & Sala, F. (2003) Evidence of a secondary grapevine domestication centre detected by SSR analysis. *Theoretical and applied genetics,* **107,** 1315-1320.

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

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters,* **16,** 1424-1435.

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

Hannah, L., Roehrdanz, P.R., Ikegami, M., Shepard, A.V., Shaw, M.R., Tabor, G., Zhi, L., Marquet, P.A. & Hijmans, R.J. (2013) Climate change, wine, and conservation. *Proceedings of the National Academy of Sciences,* **110,** 6907-6912.

Heywood, V.H. & Zohary, D. (1995) A catalogue of the wild relatives of cultivated plants native to Europe. *Flora Medit,* **5,** 375-415.

Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2010) dismo: Species distribution modeling. pp. R package version 0.7-17. Available at: http://CRAN.R-project.org/package=dismo. Last accessed 8 August 2012.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling,* **199,** 142-152.

IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, UK and New York, NY, USA.

Johnson, H. (2002) *L'atlas mondial du vin.* Flammarion, Paris.

Jones, G.V., White, M.A., Cooper, O.R. & Storchmann, K. (2005) Climate change and global wine quality. *Climatic change,* **73,** 319-343.

Kell, S., Maxted, N. & Bilz, M. (2012) 28 European Crop Wild Relative Threat Assessment: Knowledge Gained and Lessons Learnt. *Agrobiodiversity Conservation: Securing the Diversity of Crop Wild Relatives and Landraces* (ed N. Maxted), pp. 218. CABI, Wallingford, Oxfordshire; Cambridge, MA.

Kenny, G. & Harrison, P. (1992) The effects of climate variability and change on grape suitability in Europe. *Journal of Wine Research,* **3,** 163-183.

Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J. & Scott, J.K. (2011) CliMond: global high-resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution,* **3,** 53-64.

Laguna, E. (2004) American and Hybrid Grapevines (*Vitis* spp.): A new concept of invasive plants to Europe. *4th European Conference on the Conservation of Wild Plants. A workshop on the Implementation of the Global Strategy for Plant Conservation in Europe. Valencia (Spain), Poster presentation*.

Levadoux, L. (1956) *Les populations sauvages et cultivées des Vitis vinifera L.* Institut national de la recherche agronomique.

McCullagh, P. & Nelder, J.A. (1989) *Generalized Linear Models. 2nd edition.* Chapman and Hall, London.

McGovern, P.E. (2003) *Ancient wine: the search for the origins of viniculture.* Princeton University Press.

Meier, N., Pfister, C., Wanner, H. & Luterbacher, J.r. (2007) Grape harvest dates as a proxy for Swiss April to August temperature reconstructions back to AD 1480. *Geophysical Research Letters,* **34**.

Morano, L.D. & Walker, M.A. (1995) Soils and Plant Communities Associated with Three *Vitis* Species. *American Midland Naturalist,* **134,** 254-263.

Myles, S., Boyko, A.R., Owens, C.L., Brown, P.J., Grassi, F., Aradhya, M.K., Prins, B., Reynolds, A., Chia, J.-M., Ware, D., Bustamante, C.D. & Buckler, E.S. (2011) Genetic structure and domestication history of the grape. *Proceedings of the National Academy of Sciences,* **108,** 3530-3535.

Nemani, R.R., White, M.A., Cayan, D.R., Jones, G.V., Running, S.W., Coughlan, J.C. & Peterson, D.L. (2001) Asymmetric warming over coastal California and its impact on the premium wine industry. *Climate research,* **19,** 25-34.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. & Kassem, K.R. (2001) Terrestrial ecoregions of the worlds: A new map of life on Earth. *Bioscience,* **51,** 933-938.

Peterson, A.T. & Vieglais, D.A. (2001) Predicting species invasions using ecological niche modeling: New approaches from bioinformatics attack a pressing problem. *Bioscience,* **51,** 363-371.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling,* **190,** 231-259.

Poczai, P., Hyvönen, J., Taller, J., Jahnke, G. & Kocsis, L. (2012) Phylogenetic analyses of Teleki grapevine rootstocks using three chloroplast DNA markers. *Plant Molecular Biology Reporter***,** 1-16.

R (2005) *R : A language and environment for statistical computing. Version 2.5.1.* Development Core Team. R Foundation for Statistical Computing, Vienna, Austria.

Reynaud, E. (1971) *Sciences et techniques de la vigne: Culture, pathologie, defense sanitaire de la vigne.* Dunod.

Rödder, D. & Engler, J.O. (2011) Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. *Global Ecology and Biogeography,* **20,** 915-927.

Smith, C.M. (2005) *Plant resistance to arthropods.* Springer.

This, P., Lacombe, T. & Thomas, M.R. (2006) Historical origins and genetic diversity of wine grapes. *TRENDS in Genetics,* **22,** 511-519.

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

Thuiller, W., Richardson, D.M., Pysek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (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.

Verbruggen, H., Tyberghein, L., Belton, G.S., Mineur, F., Jueterbock, A., Hoarau, G., Gurgel, C.F.D. & De Clerck, O. (2013) Improving Transferability of Introduced Species' Distribution Models: New Tools to Forecast the Spread of a Highly Invasive Seaweed. *Plos one,* **8,** e68337.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. *Evolution,* **62,** 2868-2883.

White, M.A., Diffenbaugh, N., Jones, G.V., Pal, J. & Giorgi, F. (2006) Extreme heat reduces and shifts United States premium wine production in the 21st century. *Proceedings of the National Academy of Sciences,* **103,** 11217-11222.

Zecca, G., Abbott, J.R., Sun, W.-B., Spada, A., Sala, F. & Grassi, F. (2012) The timing and the mode of evolution of wild grapes (Vitis). *Molecular Phylogenetics and Evolution,* **62,** 736-747.

Zecca, G., De Mattia, F., Lovicu, G., Labra, M., Sala, F. & Grassi, F. (2010) Wild grapevine: *silvestris*, hybrids or cultivars that escaped from vineyards? Molecular evidence in Sardinia. *Plant Biology,* **12,** 558-562.

Zweig, M.H. & Campbell, G. (1993) Receiver-Operating Characteristic (Roc) Plots - a Fundamental Evaluation Tool in Clinical Medicine. *Clinical Chemistry,* **39,** 561-577.

SUPPLEMENTARY MATERIAL

https://drive.google.com/file/d/0B5ZZgCyoOEdnSThkSE1zem5DOUE/edit?usp=sharing

Fig. S1: ENM predictions for *V. berlandieri*, *V. riparia*, *V. rupestris*, *V. sylvestris* and *V. vinifera* in North America and Europe. ENM were calibrated either with 8 climate variables relevant for plant species distribution (V_{env}) or two first components of a PCA calibrated on the pooled study area (V_{pca}). Predictions were assessed under current climate conditions and future climate scenarios (A1B and A2) of two GCM (MIRO-C and CSIRO) at two reference dates (2050 and 2100) resulting 90 maps of potential distributions.

https://drive.google.com/file/d/0B5ZZgCyoOEdnYlJWR1dtMXF2RzQ/edit?usp=sharing

Figure S2: differences between ENM predictions for *V. sylvestris* and *V. vinifera* under current (a), A2 2050 (b) and A2 2100 (c) climate conditions. High values show areas more favorable to *V. sylvestris* than for viticulture. Predictions for both GCM were averaged for each reference time.

PART III

Modeling beyond the macroclimatic niche

CHAPTER 3.1

The hyper-numerical revolution to help manage biological invasions: very high resolution monitoring and modelling of alien plant distributions along a Swiss river

Patrice Descombes[1], Blaise Petitpierre[1], Eric Morard[2], Michael Berthoud[3], Antoine Guisan[1], [3] & Pascal Vittoz[1]

[1] Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland.

[2] BEB SA, Ch. des Artisans 6, CH-1860 Aigle, Switzerland.

[3] Institute of Earth Sciences, University of Lausanne, CH-1015 Lausanne, Switzerland.

This paper is in review in *Biological Invasions*

My contribution to the project: This is a master project I co-supervised with PV. I supervised the preliminary studies, initiated the master student to fieldwork, and designed the statistical analyses. Both first authors contributed equally.

ABSTRACT

Monitoring and species distribution models (SDMs) are increasingly used to support conservation planning but rarely projected at a fine scale for conservation management. In this study, we compared population distribution and size of five invasive plant species along an 18 km alluvial system in Switzerland, over a period of 11 years. We compared exhaustive inventories of past (2001) to current (2012) populations showing a massive increase in invaded areas in eleven years. *Impatiens glandulifera* and *Reynoutria japonica* were the species with the largest increase in population number and size. The ecological preferences and potential distribution of each species was modelled through SDMs at 1 m resolution, using environmental variables referring to topography, disturbances, dispersal, soil texture and light availability. From these models, we projected species potential distribution and derived both potential and actual range filling indices, which together indicate the habitat still available for species colonisation. SDMs successfully depicted the niches at very high resolution. Some of the important predictors (e.g., canopy density, river sinuosity) could not have been measured at a coarser resolution. These models predicted large, empty areas still suitable for each species suggesting that the increase of the population number and size observed in this study is expected to continue in the coming years. The very high resolution SDMs reveal precisely areas potentially available for future invasions, and can therefore be used to prioritise eradication efforts in already invaded areas and controls in areas at high risk of invasion.

INTRODUCTION

The invasion of natural communities by non-native species is one of the major causes of biodiversity loss added to habitat decline (Keller *et al.* 2012). Invasive plant species (IPS) are able to alter nutrient cycling (Ehrenfeld 2003) and to outcompete the native flora by shading out species (Lake and Leishman 2004). Although mostly associated with human activity around urban areas (Hulme 2003), IPS are also known to be very abundant along rivers where water flow acts as a dispersal vector, particularly when flooding occurs (Tickner *et al.* 2001). Such disturbances by water creates gaps in the forbs cover, providing suitable habitats and favourable nutrient conditions for the establishment of seedling IPS (Renöfält *et al.* 2005; Maskell *et al.* 2006; Miller and Matlack 2010).

Human activity (e.g., river containment) during the 19th and 20th centuries drastically decreased the number of rivers presenting a natural dynamic (alluvial areas) in Switzerland (Lachat *et al.* 2010). Because these habitats are also considered to be hotspots of native species diversity (Ward *et al.* 2002), and at the same time are at a high risk of invasion by IPS (i.e., represent conflict areas; Vicente *et al.* 2011), there is an urgent need to implement conservation actions to limit the impact of IPS in these rare and fragile ecosystems.

Biological invasions are characterised by a slow initial spread (lag-phase) where species occur only in a few localities, followed by a fast spread (exponential phase) and a third

phase of decelerating expansion (Pysek and Hulme 2005). However, invasion processes comprise both local population expansion and colonisation of new areas on larger scales (Pysek and Hulme 2005). Thus, the scale at which the biological invasion is assessed can lead to different interpretations. The growth of local populations does not necessarily cause range expansion, whereas range expansion does not necessarily implies the growth of local populations. Studies reporting colonisation of new areas by IPS are numerous (e.g., Pysek and Prach 1993; Pysek and Hulme 2005), but examples of population monitoring at the local scale are much rarer (e.g., Kasperek 2004) despite their crucial importance for understanding colonisation processes (Pysek and Hulme 2005).

Furthermore, prevention of biological invasion at the earliest stage is more effective than attempts at management of well-established infestations (Leung *et al.* 2002). Species distribution models (SDMs; Guisan and Thuiller 2005) represent potentially useful tools in this regard (Guisan *et al.* 2013). They relate environmental characteristics (e.g., climatic and topographic variables) to the current geographical distribution of species in terms of presence/absence or abundance to fit species' realized niches and predict the distribution of suitable habitats. Additionally, SDMs can also prove useful in conservation by detecting and ranking environmental variables affecting the distribution and fitness of species and by helping managers visualise their invasive potential (Guisan and Thuiller 2005).

During the past decade, most SDMs studies investigated the distribution of invasive species at coarse resolution (≥ 100 m) across large areas, corresponding to potential range expansion. However, coarse resolution mapping tends to overestimate the potential distribution of invasive species and to describe inadequately the spatial structure of invasions at the local population scale (Hulme 2003). Thus, fine resolution SDMs are required for conservation efforts, planning and management, which are performed at a local scale and need precise knowledge of the species' ecology (Heinänen *et al.* 2012).

By comparing past population distributions to current exhaustive inventories, we quantified changes in population number and size for four invasive species (*Reynoutria japonica*, *Impatiens glandulifera*, *Helianthus tuberosus* and *Buddleja davidii)* along 18 km of a Swiss river. We then fitted the niches of the aforementioned species and *Prunus laurocerasus* using existing data within very high resolution (1 m) SDMs. Accordingly, the aims of the study were (1) to quantify the changes in population size of invasive plant species along the river between 2001 and 2012, (2) to identify the environmental variables explaining species presence and abundance at very high resolution, and (3) to identify the areas potentially vulnerable to future invasions. Finally, the findings are discussed in the context of management of protected areas with regard to IPS.

METHODS

Study area

The study area is located in western Switzerland along an 18 km section of the Venoge River (46.57 N, 6.53 E; Fig. 1). This alluvial area, ranging from 370 to 450 m a.s.l., is located in the colline vegetation belt, naturally composed of deciduous forests dominated by *Quercus robur* and *Fraxinus excelsior*. The Venoge undergoes heavy flooding mainly in winter, regularly modifying the river's course and maintaining the diversity of alluvial vegetation, a situation rarely sustained on the Swiss Plateau. This conservation value is recognised by the National Inventory of Natural Alluvial Areas (Swiss Federal Office for the Environment FOEN, http://www.bafu.admin.ch), which includes two protected areas within the study area. To focus the inventory only on the invasive species in an alluvial dynamic context, the study area was bounded either by a width of 30 m on each side of the river or an elevation of 2 m above the mean water level with a 1 m resolution digital elevation model (DEM) (MNT-MO obtained by LiDAR (Light Detection And Ranging); Office of Land Information OIT, www.vd.ch/oit) using ArcGIS 9.3 software (ESRI, Redlands, CA, USA, 2008).

Population inventory

Between July and September 2012, we inventoried the entire population of five IPS along the study area: *Reynoutria japonica* Houtt., *Impatiens glandulifera* Royle, *Helianthus tuberosus* L., *Buddleja davidii* Franch. and *Prunus laurocerasus* L. (see Table 1 for a summary of their basic life history characteristics). *R. x bohemica* Chrtek & Chrtková (hybrid of *R. japonica* and *R. sachalinensis* (F. Schmidt) Nakai) was included in *R. japonica* as individuals exhibited intermediate traits, making species identification questionable. These five species are listed as invasive on the "Black list" and "Watch list" of the Swiss Federal Office for the Environment (http://www.infoflora.ch). A population was defined as a set of individuals closer than five meters from each other. For each population, the precise location was recorded with a Trimble GeoXH handheld GPS connected to a Tornado external antenna (sub-metric precision < 0.5 m); the area occupied was estimated in m^2 and the number of individuals was precisely counted for small populations or estimated for large populations (> 300 individuals). An individual was defined as a single stem for *H. tuberosus* and *I. glandulifera* and as a bush for *B. davidii* and *P. laurocerasus*. Because *R. japonica* forms dense populations consisting of groups of stems, it would have been laborious and unreliable to make a count of all the stems; an individual was defined as a group of stems (closer than 15 cm to one another).

Table 1. Characteristics of the investigated species in their invaded range.

	R. japonica	***I. glandulifera***	***B. davidii***	***H. tuberosus***	***P. laurocerasus***
Area of origin	East Asia	Himalaya	China	North America	Armenia
Life form (Raunkiær, 1934)	Hemicryptophyte	Therophyte	Phanerophyte	Geophyte	Phanerophyte
Morphology	Robust stems rhizomes	Robust stems	Woody shrub	Stems tubers	Woody shrub evergreen
Height (max)	3 m	2 m	5 m	3 m	8 m
Habitat:					
- canopy	Light	Light, shade	Light	Light	Shade
- Soil type	various, disturbed	various, disturbed, (wet/rich)	various, disturbed	various, (wet/rich)	organic
Reproduction	Vegetative	Seeds	Seeds, vegetative	Vegetative	Seeds, vegetative
Dispersal	Water, human	Water, human	Water, wind	Water, human	Birds, (water)

Sources: Swanton *et al.*, 1992; Beerling & Perrins, 1993; Beerling *et al.*, 1994; Tallent-Halsell & Watt, 2009; National data center on the Swiss flora (http://www.infoflora.ch)

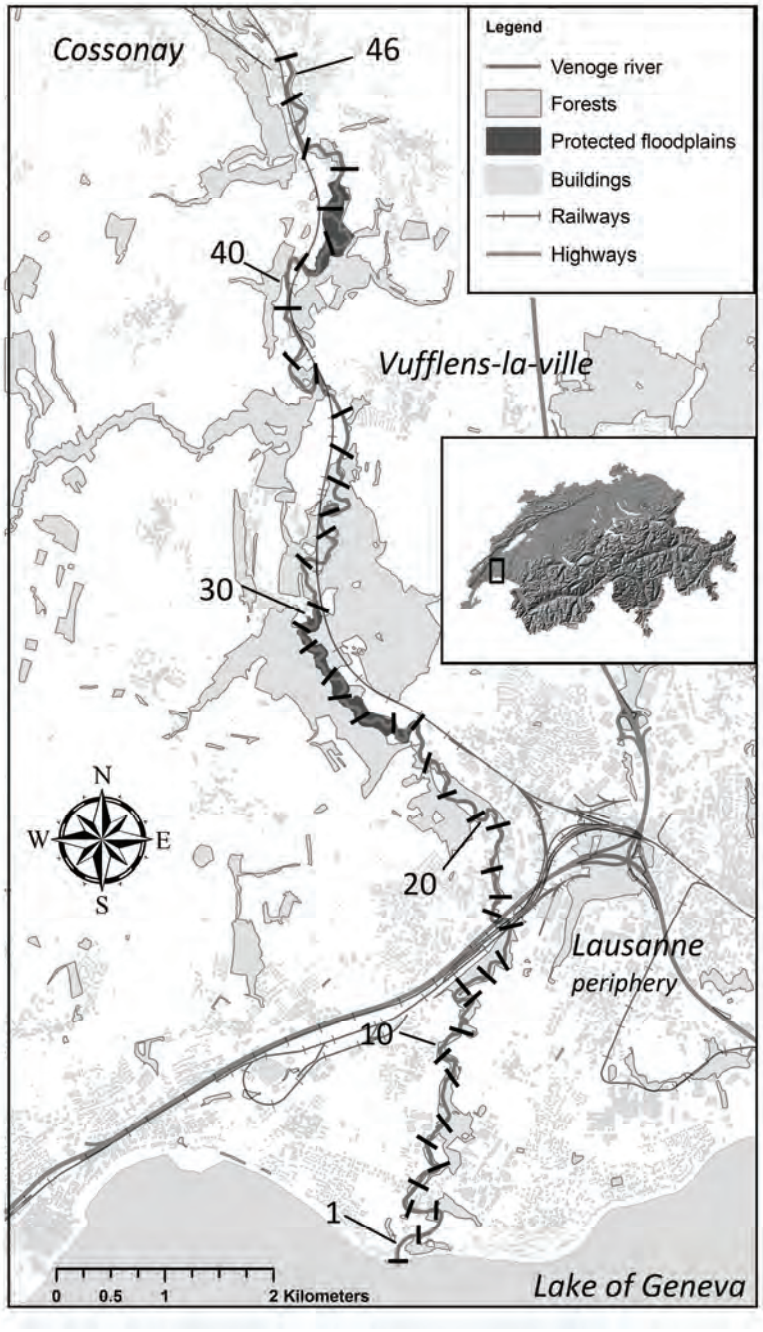


Figure 1: Location of the study area in western Switzerland. Study area is limited by the potentially flooded zones along 18 km of the Venoge River, split into 46 sectors approximately 400 m in length.

Historical data

An inventory performed during the summer of 2001 along the same study area already reported the population distribution of *R. japonica*, *I. glandulifera*, *H. tuberosus* and *B. davidii* (Morard 2001, unpublished data). For this first inventory, the river was divided into 46 sectors, each approximately 400 m long (Fig. 1). Populations of *R. japonica* and *B. davidii* were inventoried and classified into five size classes according to the number of individuals (class 1, 1-4 individuals; 2, 5-9; 3, 10-49; 4, 50-99; 5, >100). In addition, the areas (m^2) occupied by the populations were estimated. *H. tuberosus* and *I. glandulifera* individuals were counted for each sector.

We compared the density (individuals per linear metre of river) of the four IPS within each sector. *R. japonica* and *B. davidii* densities for 2001 were extrapolated from the population size (classes 1 to 5) following three conversions: minimum, median and maximum numbers of individuals in size classes. These conversions were performed to define the possible range of changes of density between 2001 and 2012. To estimate the number of individuals in population of class 5 reported in 2001, we estimated the mean density (individual/m^2) in populations of class 4 sampled in 2012 and extrapolated to the area estimated for populations of class 5 sampled in 2001. For three species (*R. japonica*, *I. glandulifera* and *B. davidii*), some populations observed in 2001 could be relocated in 2012 on the basis of the original map (established without GPS) or photograph archives, and population size changes could be assessed at the population level. A non-parametric paired test (Wilcoxon signed rank test) was performed to assess differences between 2001 and 2012 for sector densities and population sizes.

Environmental variables

We selected 17 environmental variables that could explain the species distribution at fine resolution (Table 2). All variables refer to topography, disturbances, dispersal, soil texture or light availability and were derived from GIS layers using ArcGIS 9.3 software (ESRI, Redlands, CA, USA, 2008) with 1 m resolution, except soil texture, canopy density and containment influence, which were measured in the field. All the variables were sampled at each population site. Additionally, sixty sites were selected using a random-stratified sampling design based on distance to the river. These sites, devoid of IPS, were similarly distributed to the river as the IPS populations in order to characterise the average conditions (or “background data”) in the study area.

Soil texture was estimated visually (in percentage) at a depth of 5 cm for the following four classes: organic matter, fine mineral sediments (< 2 mm), small stones (2-75 mm) and large stones (> 75 mm). These visual estimations were then corrected using a standardised measured reference (see Appendix S1 in Supporting Information). Topography variables, including slope (degree) and curvature (degree of convexity/concavity), were derived from the DEM at 1 m resolution using the “slope” and “curvature” functions in ArcGIS. Variables related to dispersal included distance to the river, distance to major roads, distance to forest roads, distance to railway lines, density of habitation (including farms; %) and density of industrial areas (%). Disturbance variables included containment presence along the river (binary yes/no), protection status (binary yes/no), river sinuosity (index) and river curvature (index). Containment influence was derived from field mapping, with perpendicular influence along the corresponding river side, and protection status corresponded to the areas delimited by the Swiss Inventory of Natural Alluvial Areas (FOEN; http://www.bafu.admin.ch). These two variables are proxies of the environmental quality. Sinuosity corresponds to the ratio between the 100 m separating two points along the watercourse and the shortest straight path connecting them. A high sinuosity indicates that the river presents a tortuous watercourse where open

areas and sedimentation processes are favoured (Tickner *et al.* 2001; Ward *et al.* 2002). Curvature corresponds to the radius of the river curves. A positive value of river curvature corresponds to the internal shore of the curves where sedimentation is important (i.e., inside meander), whereas a negative value corresponds to the external shore of the curves where erosion is high (i.e., outside meander). Both sinuosity and curvature were calculated for each meter along both riversides. Note that distance to the river can also be seen as a proxy of river disturbance. We measured the canopy density with two techniques. First, canopy closure was measured directly in the field at population and background sites (CC_{Field}) during the field sampling in 2012. Second, canopy cover was obtained by remote sensing, using GIS layers (sampled in 2001-2002), to estimate light availability for all pixels of the study area and included in model spatial projections (CC_{LiDAR}). See Appendix S1 for methodological details about indices estimates.

Quantifying ecological niche separation between species

Because some variables were not quantitative, we performed a mixed multivariate analysis (MMA) following the Hill and Smith method (1976), a form of non-parametric principal component analysis, to quantify the separation between the ecological niches of the five species. MMA was realized on the populations observed in 2012 and the sites randomly selected without IPS, explained by the 17 environmental variables (Table 2) in the R software (R Development Core Team 2011) using the package ade4 (Dray and Dufour 2007).

Environmental niche modelling

To determine the ecological preferences of the five IPS, we built SDMs by relating all the presence-absence or abundance (i.e., number of individuals per population) observations recorded in 2012 with the 17 environmental variables (Table 2) using custom code in R. For this purpose, we ran an ENSEMBLE approach (Araújo and New 2007) by averaging the results of three commonly used statistical techniques: a generalized linear model (GLM; McCullagh and Nelder 1989), a gradient boosting model (GBM; Ridgeway 1999; Friedman *et al.* 2000) and MAXENT (Phillips *et al.* 2006; Elith *et al.* 2011). GBM and MAXENT were performed using the package dismo (Hijmans *et al.* 2011). Only GLM and GBM were used to model abundance, as MAXENT does not include an appropriate distribution for count data. Relevant variables were *a priori* selected following a forward-backward stepwise procedure using the StepAIC function (Venables and Ripley 2002) and by a k-fold cross-validation function (Elith *et al.* 2008) in GLM and GBM, respectively. The importance of variables was assessed by estimating the effect of variable randomization on species suitability predictions following Thuiller *et al.* (2010).

Table 2. Environmental variables selected for the niche modeling (description as complete as possible of the species ecological niche) and for the model projections (potentially suitable areas for species colonization), derived from GIS layers or measured in the field (*). Canopy density was measured in the field (CCField) and derived by remote sensing as a proxy of light availability (CCLiDAR). Model projection is restricted to the variables available for the whole study area, what excludes soil variables.

Environmental variables	Abbreviation	Units	Niche modelling	Model projections
Distance to the river	River	meter	yes	yes
Distance to major roads	M. roads	meter	yes	yes
Distance to forest roads	F. roads	meter	yes	yes
Distance to railway lines	Railway	meter	yes	yes
Density of industry	Industry	%	yes	yes
Density of habitation	Habitation	%	yes	yes
Slope	Slope	degree	yes	yes
Curvature	Curvature	degree	yes	yes
Containment*	Contain.	binary (yes/no)	yes	yes
Protection status	P. status	binary (yes/no)	yes	yes
River sinuosity	River sin.	index	yes	yes
River curvature	River curv.	index	yes	yes
Canopy density*	Canopy	%	yes (CC_{Field})	yes (CC_{LiDAR})
Organic matter*	Organic	%	yes	no
Fine sediments (< 2 mm)*	Sediments	%	yes	no
Small stones (2-75 mm)*	S. stones	%	yes	no
Large stones (> 75 mm)*	L. stones	%	yes	no

To assess model predictions, we used the area under the ROC-plot curve (AUC; Fielding & Bell 1997) and the True Skill Statistics (TSS; Allouche *et al.* 2006), which both evaluate the ability of the model to discriminate presences from absences. AUC varies between 0 (counter prediction) and 1 (perfect prediction), 0.5 meaning random predictions. TSS is scaled between -1 and 1, 0 meaning random predictions. However, as biological invasions are ongoing processes, absences may not be representative of true absences and may thus bias the evaluation (Václavík and Meentemeyer 2009). We thus added two presence-only evaluators: the Boyce index (Boyce; Hirzel *et al.* 2006) and the sensitivity calculated on predictions binarised with the threshold providing the best TSS. The Boyce index measures how much model predictions differ from the random distribution of the observed presences across the range of prediction values and is analogous to a correlation value, varying between -1 (counter prediction) and 1 (perfect prediction), 0 meaning random predictions. The sensitivity is a threshold-dependent evaluator corresponding to the rate of presences correctly classified by the model. For the abundance models, we used Spearman's rank correlation (SC).

Data were split randomly into two partitions: 70 % was used for model calibration, and the remaining 30 % was used for model evaluation. This procedure was replicated 10 times, and evaluators and variable importance were averaged for all models and replicates.

To project models and to visualize potential species distribution in the whole study area, we built SDMs using the 13 environmental variables available for the whole study area (Table 2) following the same protocol as described above. Then, for each species we calculated a potential range filling index (PRF) and an actual range filling index (ARF) per sector i:

$PRF = h_i / h_{max}$

$ARF = (O_i / L_i) / (O / L)_{max}$

where h corresponds to the average suitability per sector, O to the number of individuals in the sector i and L to the length of the sector i. Therefore, a high PRF value indicates that the sector contains a large area of suitable habitats that is more susceptible to invasion than other sectors. If coupled with low ARF, it indicates that the sector contains a large area of suitable habitat but few invasive individuals. However, if coupled with high ARF, it indicates that the suitable habitat in the sector has already been largely invaded. Thus, sectors with a low difference between PRF and ARF values are closer to IPS saturation than other sectors.

RESULTS

General observations

A total of 983 populations of invasive plants were recorded in the study area during summer 2012, including populations of *R. japonica* (N = 381), *I. glandulifera* (N = 313), *B. davidii* (N = 147), *P. laurocerasus* (N = 88) and *H. tuberosus* (N = 54). *I. glandulifera* presented the highest proportion of large populations (37.06 % of class 5, > 100 individuals), followed by *R. japonica* (12.07 %), *H. tuberosus* (11.11 %) and *P. laurocerasus* (2.27 %), whereas *B. davidii* had none. Species' densities (number of individuals per metre of sector) was the highest, on average, for *I. glandulifera* (mean ± SD, 4.25 ± 6.12), followed by *R. japonica* (1.11 ± 1.35), *H. tuberosus* (0.13 ± 0.23), *P. laurocerasus* (0.07 ± 0.17) and *B. davidii* (0.03 ± 0.06). The total area occupied was 29'804 m^2 for *I. glandulifera*, 22'775 m^2 for *R. japonica*, 6'545 m^2 for *B. davidii*, 5'464 m^2 for *P. laurocerasus* and 1'089 m^2 for *H. tuberosus*.

Changes in invasive species densities between 2001 and 2012

The changes in sector densities observed between 2001 and 2012 (Fig. 2) were significant for *I. glandulifera* (Wilcoxon signed rank test, P-value < 0.001, N = 37), with an average increase of 1122%, and under the three conversions for *R. japonica* (P-value < 0.001, $\alpha_{Bonferroni}$ = 0.0083, N = 44), with an increase of 565 % (using the median value). For *B. davidii*, the difference was significant only with the minimum value (P-value < 0.001, $\alpha_{Bonferroni}$ = 0.0083, N = 35), with an average increase of 85 % (median value), whereas there was no significant variation for *H. tuberosus* (P-value = 0.58, N = 28). The changes

in size of the 2001 populations relocated in 2012 (see Fig. S1) were significant for *I. glandulifera* (P-value = 0.027, N = 10) and under the three conversions for *R. japonica* (P-value < 0.001, $\alpha_{Bonferroni}$ = 0.0083, N = 66), but none of the conversions was significant for *B. davidii* (P-value > 0.049, $\alpha_{Bonferroni}$ = 0.0083, N = 23). *P. laurocerasus* was not inventoried in 2001, whereas the species was too rare to be considered as a problem in regional forests.

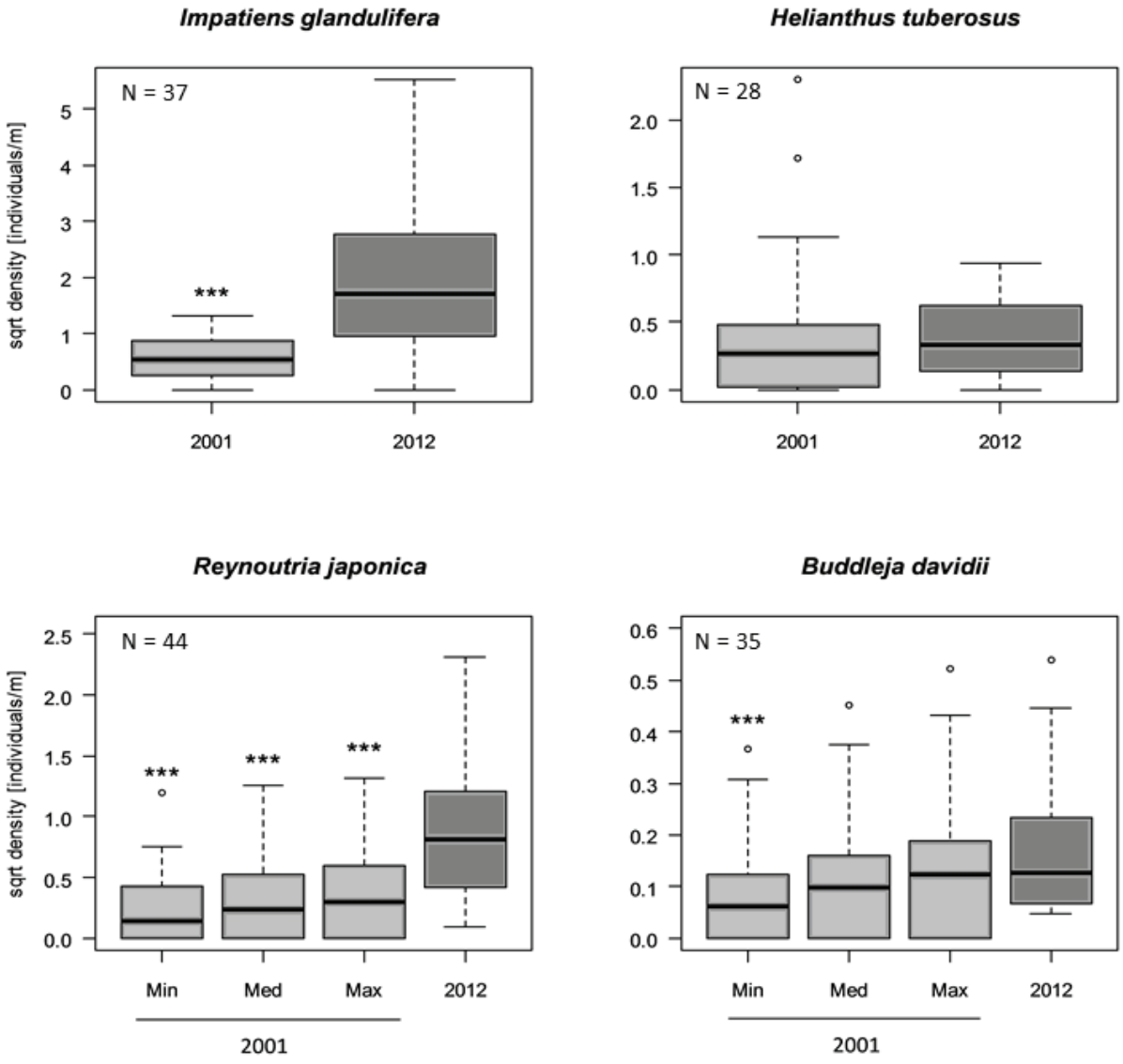


Figure 2: Evolution of the IPS density (individuals/m) along 46 sectors of the Venoge River between 2001 and 2012. Sectors with no observations were removed from the analysis. *R. japonica* and *B. davidii* densities for 2001 were extracted from the population size classes (classes 1 to 5) following three conversions: minimum (Min), median (Med) and maximum (Max) number of individuals in size classes. Two outliers of *H. tuberosus* in 2001 (i.e. 2.95 and 5.27) were not represented. The intermediate line is the median and the limits of the boxes are the 1st and 3rd quartiles. *I. glandulifera* and *R. japonica* are the species presenting a significant increase of their density between 2001 and 2012 (***, p < 0.001). Note that the scales are different between species.

Niche separation between species

The three first axes of the mixed principal component analysis (MMA) explained 37.06 % of the total variance (axis 1: 15.09 %; axis 2: 12.77 %; axis 3: 9.20 %; Fig. 3). The position of the niche centroids of the species formed two main groups composed of (1) *R. japonica*, *I. glandulifera*, *H. tuberosus* and *B. davidii*, and (2) *P. laurocerasus*. Distance between centroids indicated that the niche of *P. laurocerasus* was well separated from the other species (see Table S1). However, *R. japonica*, *I. glandulifera*, *H. tuberosus* and *B. davidii* shared a quite similar niche in the study area (Fig. 3, see Fig. S2 and Table S1).

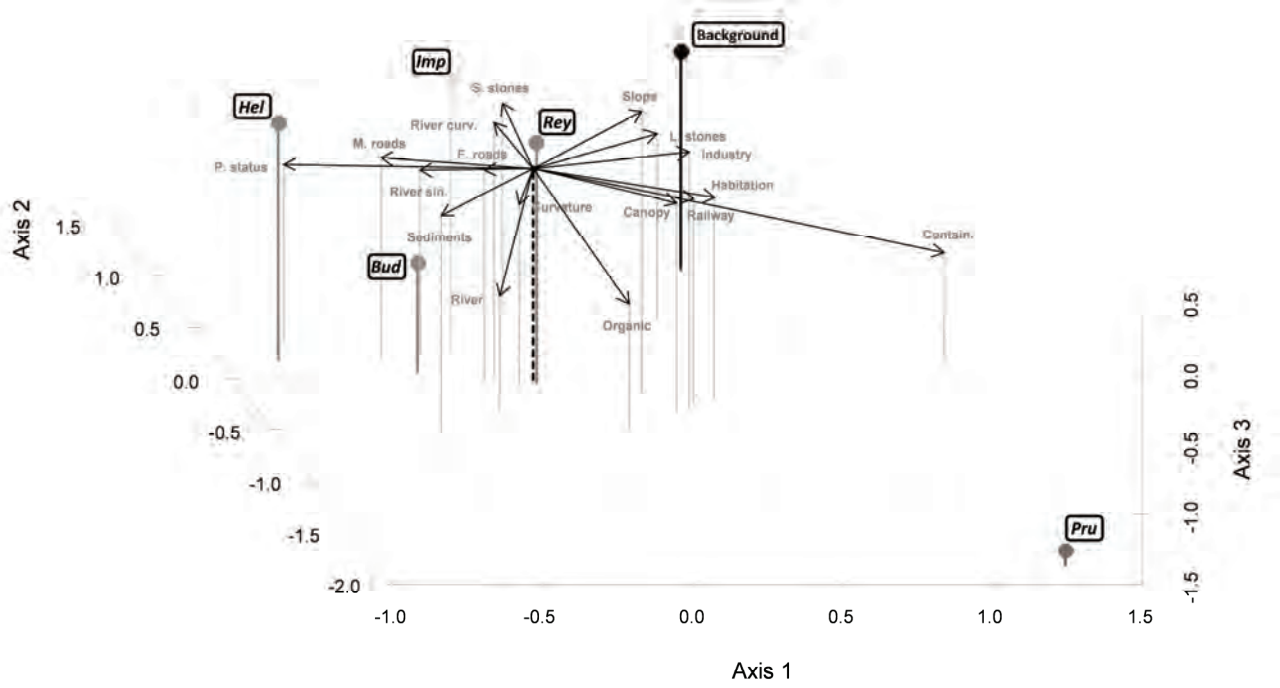


Figure 3: Mixed multivariate analysis (MMA) built with 17 environmental variables and showing the environmental distance between the niche centroids of R. japonica (Rey), I. glandulifera (Imp), B. davidii (Bud), H. tuberosus (Hel), P. laurocerasus (Pru) and the random points (Background). The three first axes explained 37.06 % of the data variation (axis 1: 15.09 %; axis 2: 12.77 %; axis 3: 9.20 %). The length of the vectors represents the magnitude of the correlation between the variables and the axes. Variables abbreviations are given in Table 2. The niche centroid of P. laurocerasus is well separated from the other species.

Models performance

All the models showed reliable predictions (Table 3). For presence-absence models including field sampled parameters, AUC ranged between 0.749 and 0.937, TSS between 0.478 and 0.812, sensitivity between 0.754 and 0.877, and the Boyce index between 0.574 and 0.694. Thus, occurrence models vary from “useful” for the weakest fitted species (*R. japonica*) to excellent for the best fitted species (*P. laurocerasus*). Abundance models significantly correlated with observed abundances, with SC between 0.330 and 0.662 (Spearman’s rho correlation test: p-value < 0.001). The models calibrated with the 13

variables available for the whole area showed similar evaluations (Table 4), thus validating the spatial projections of models.

Table 3: Performance of the species occurrence and abundance models under several evaluators as the average of ten replicates. 17 environmental variables were used, including field measurements. Occurrence models were built with three algorithms (GLM, GBM and MAXENT) under binomial distribution and were evaluated with two presence-absence evaluators (AUC, TSS) and two presence-only evaluators (Boyce, Sensitivity). Abundance models were built with two algorithms (GLM and GBM) under Poisson distribution and were evaluated with a Spearman's rank correlation (SC).

Model type	Evaluators	Species				
		R. japonica	*I. glandulifera*	*B. davidii*	*H. tuberosus*	*P. laurocerasus*
Occurrence (GLM, GBM, MAXENT)	AUC	0.749 ± 0.029	0.812 ± 0.030	0.852 ± 0.036	0.815 ± 0.025	0.937 ± 0.033
	TSS	0.478 ± 0.045	0.590 ± 0.050	0.628 ± 0.069	0.612 ± 0.033	0.812 ± 0.059
	Boyce	0.574 ± 0.136	0.694 ± 0.103	0.678 ± 0.129	0.625 ± 0.124	0.614 ± 0.210
	Sensitivity	0.754 ± 0.053	0.871 ± 0.042	0.808 ± 0.046	0.850 ± 0.062	0.877 ± 0.051
Abundance (GLM, GBM)	SC	0.330 ± 0.071	0.422 ± 0.050	0.414 ± 0.082	0.386 ± 0.056	0.662 ± 0.088

Table 4: Performance of the species occurrence and abundance models, calibrated with 13 variables available on GIS layers for the whole study area, under several evaluators as the average of ten replicates. Models were built and evaluated as in Table 3. An example of model projection is given in Fig. S6.

Model type	Evaluators	Species				
		R. japonica	*I. glandulifera*	*B. davidii*	*H. tuberosus*	*P. laurocerasus*
Occurrence (GLM, GBM, MAXENT)	AUC	0.760 ± 0.035	0.754 ± 0.040	0.869 ± 0.029	0.760 ± 0.047	0.932 ± 0.038
	TSS	0.451 ± 0.053	0.492 ± 0.051	0.650 ± 0.058	0.511 ± 0.061	0.807 ± 0.069
	Boyce	0.650 ± 0.158	0.616 ± 0.105	0.731 ± 0.075	0.593 ± 0.122	0.634 ± 0.153
	Sensitivity	0.725 ± 0.047	0.859 ± 0.044	0.824 ± 0.042	0.829 ± 0.093	0.883 ± 0.044
Abundance (GLM, GBM)	SC	0.298 ± 0.074	0.390 ± 0.058	0.438 ± 0.089	0.329 ± 0.086	0.552 ± 0.138

Importance of environmental predictors

The importance of the 17 environmental predictors used varied between species and models (Fig. 4). The predictors that best explained the species distribution for the occurrence models were distance to the river, fine mineral sediments, organic matter and canopy density (Fig. 4a). *H. tuberosus*, *I. glandulifera* and *R. japonica* were mainly found close to the river (distance < 15 m), in sites where mineral sediments were abundant, whereas *B. davidii* and *P. laurocerasus* occurrences were less influenced by proximity to the river (see Fig. S3). *P. laurocerasus* preferred organically rich forest soils and was the only species that responded positively to canopy density.

The predictors that best explained the species' abundances were distance to the river, canopy density, fine mineral sediments and organic matter (Fig. 4b). *H. tuberosus*, *I. glandulifera* and *R. japonica* were more abundant at a distance between 8 and 15 m from the river, in sites rich in mineral sediments, whereas *B. davidii* and *P. laurocerasus* were more abundant between 15 and 25 m (see Fig. S4). Species abundances were higher in open places (low canopy density), except *P. laurocerasus*, which preferred shady places with a soil rich in organic matter. In addition, *I. glandulifera*, *B. davidii* and *R. japonica* abundance were more sensitive to decreases in light availability.

The importance of predictors was very similar when only the 13 available variables for the whole area were used, with a large dominance of the distance to the river and of the canopy density in occurrence and abundance models (see Fig. S5). Hence, projections obtained with these models correspond to the ecological niche obtained from the complete models, except the absence of data on soil characteristics.

Predictions of invasion risks

The PRF calculated for each sector on the basis of the species' occurrence projections indicated the sectors that were the most susceptible to colonisation by the species (Fig. 5). By coupling this index with the 2012 ARF, we measured the relative saturation for each sector, which showed large areas at risk of future colonisation by each species. The models predicted less favourable habitats for I. glandulifera than for the other species, as light availability acted as a limiting factor (Fig. 5 and S6). The distribution pattern of R. japonica, *I. glandulifera*, *H. tuberosus* and *B. davidii* observed in 2012 indicates that these species were found relatively often together in the same sectors. By comparing the 2001 ARF to the 2012 ARF, we measured the relative increase in species' densities. Globally, all sectors showed an increase of their invasive species' densities, except *H. tuberosus*, which showed an important decrease in some sectors (i.e., sector 26, 27 and 31). Interestingly, these sectors were also the ones presenting the highest increases of *I. glandulifera* and *R. japonica*.

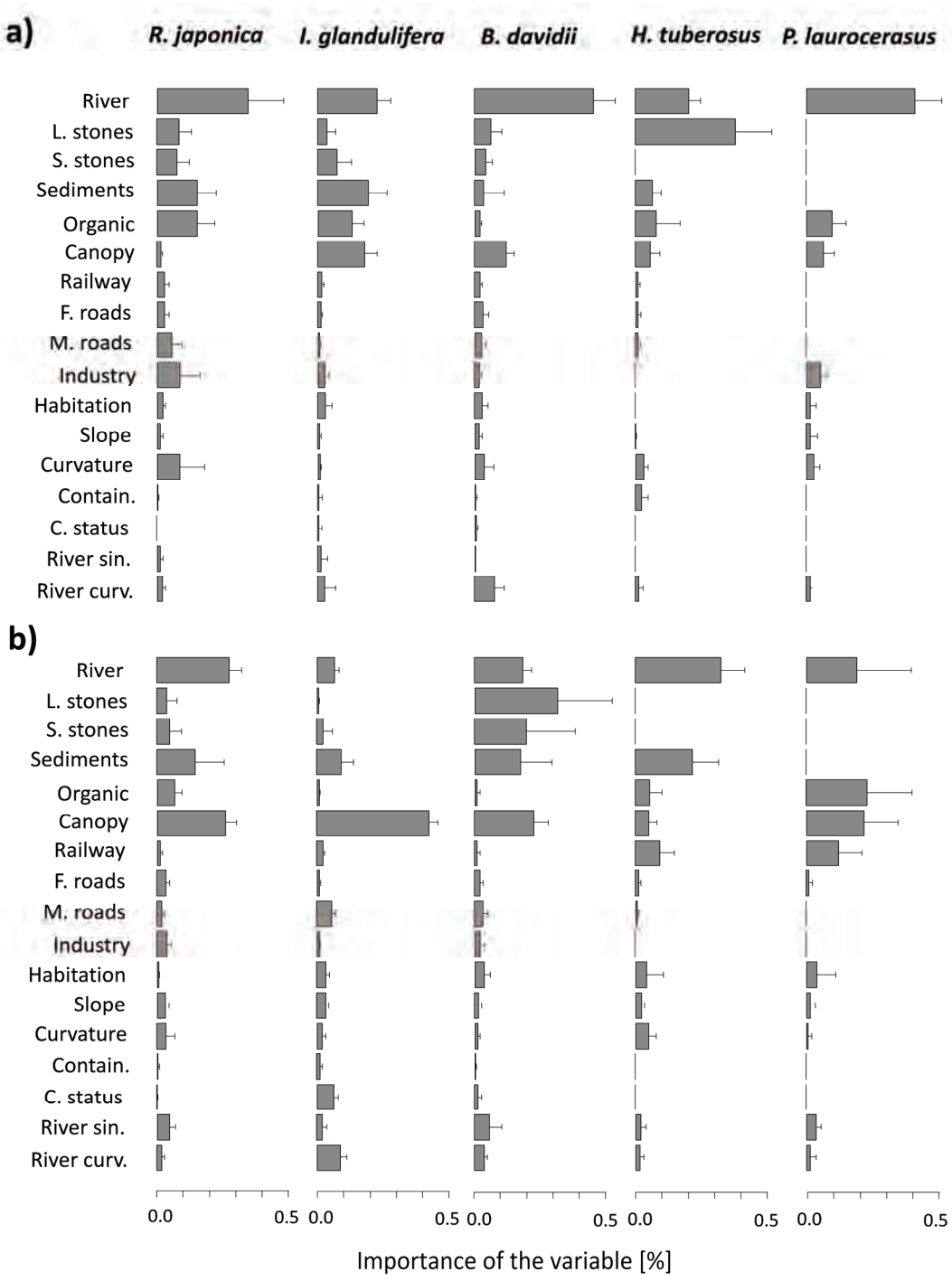


Figure 4. Importance (assessed by permutation, see methods) of the variable in the niche models for species (a) occurrence and (b) abundance models. Variable abbreviations are given in Table 2.

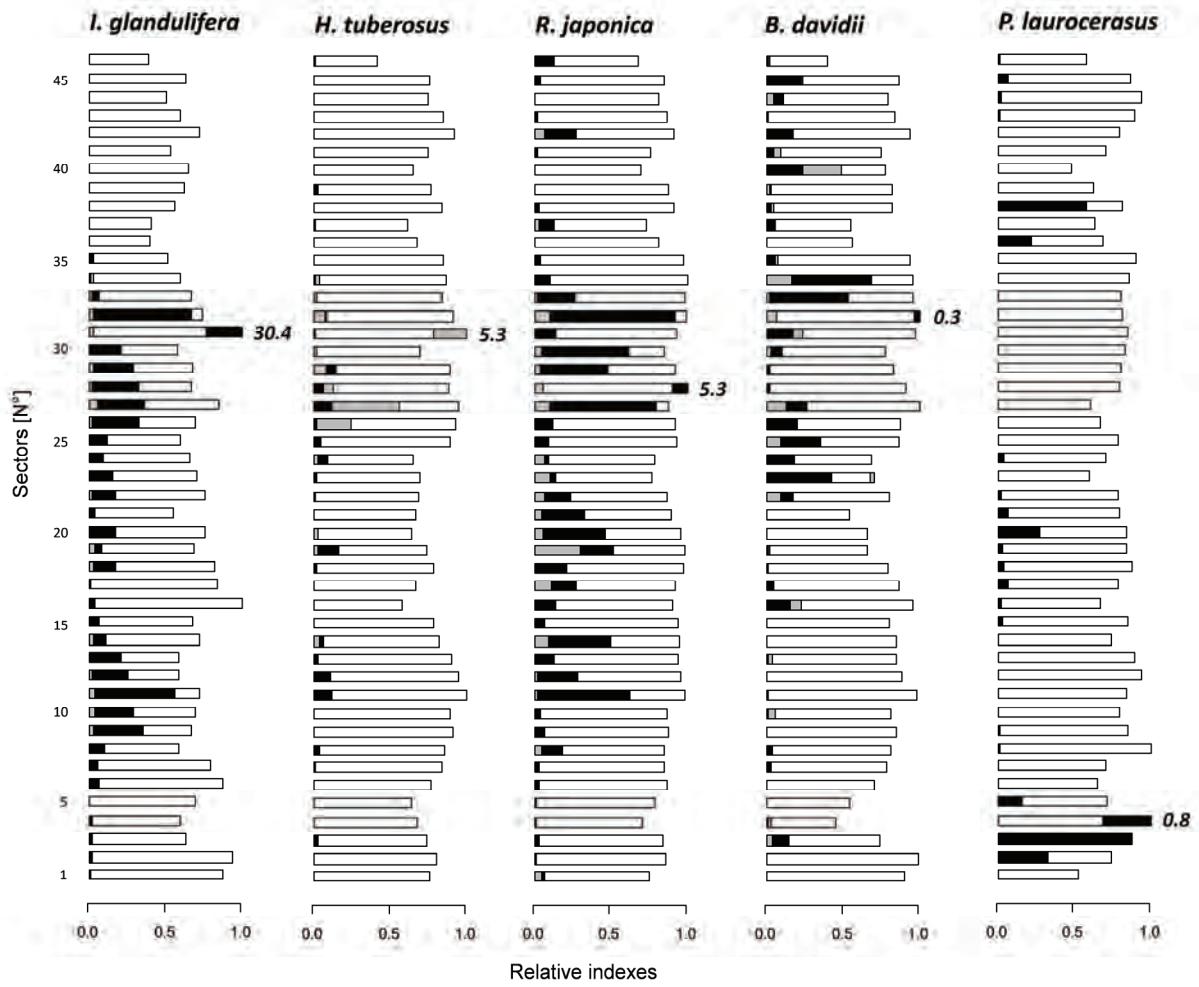


Figure 5. Potential range filling index (PRF, white bars) and actual range filling index (ARF) for 2001 (grey bars) and 2012 (black bars) along 46 sectors of the Venoge River from the Lake of Geneva (sector 1) to the village of Cossonay (sector 46). The PRF values correspond to the relative amount of suitable habitat present in each sector, and the ARF values the relative species density. Values next to the bars correspond to the absolute density (individuals/m) of the most invaded sector between 2001 and 2012.

DISCUSSION

Our study showed an overall massive increase of invasive plant species along the Venoge River over a period of 11 years. We were able to successfully model their potential habitat at a very high resolution using SDMs and project them across the entire study area. From these results, we measured potential range filling (PRF) coupled with actual range filling (ARF) per sector, which together showed that large areas of suitable habitat for each species remained unoccupied, despite some of the species already being abundant in the study area. These models are thus useful for prioritising management of IPS.

Species environmental preferences

The abiotic environmental preferences highlighted by our models are in general consistent with literature on the species' ecologies. Flood disturbances disperse IPS diaspores (Pysek and Prach 1993; Tickner *et al.* 2001), erode river margins, deposit mineral sediments and create openings in the forbs or tree cover, microhabitats favourable to seed germination (Renöfält *et al.* 2005; Maskell *et al.* 2006; Miller and Matlack 2010). The disturbances therefore create good opportunities for IPS establishment by increasing light availability, soil resources and reducing species competition (Davis *et al.* 2000; Miller and Matlack 2010), which explains why IPS are widespread along river corridors (Renöfält *et al.* 2005). Consistent with these observations, our models showed that *R. japonica*, *I. glandulifera*, *B. davidii* and *H. tuberosus* prefer sites with high light availability in sinuous river sectors that create new primary areas by meander expansion and diaspore deposition with sediments. However, *B. davidii* is more tolerant to drought and prefers stonier soils than *R. japonica*, *I. glandulifera* and *H. tuberosus*, which are found on moist sediment and nutrient rich soils (Wyse *et al.* 1986; Barney *et al.* 2006; Pysek *et al.* 2012). Interestingly, light availability appears to have a lower influence on presence-absence than on the abundance of *R. japonica*, *I. glandulifera* and *B. davidii* (Fig. 4), supporting that these species can grow in areas with variable light regime but find optimal growth conditions in open areas. Thus, the possible colonisation of shaded habitats outside floodplains is not excluded, as observed regularly for *R. japonica* and recently for *I. glandulifera* (Pysek *et al.* 2012). Although there is high suitability all along the study area for *I. glandulifera*, the upstream part of the river (sector 35 to 46) is less invaded than the downstream sectors (1 to 35), suggesting an initial introduction near the village of Vufflens-la-ville (sector 35; see ARF values, Fig. 5) and later downstream dispersal, underlining the possible importance of river dispersal (Pysek and Prach 1993). Only the *P. laurocerasus* niche can be clearly distinguished from the other IPS by mostly colonising shaded, understory conditions. It is a highly shade-tolerant species, a uncommon characteristic among native woody species (Landolt *et al.* 2010), that depends on dispersal by birds (Richardson *et al.* 2000), especially *Turdus merula*, a very common bird in Swiss forests (Glutz Von Blotzheim 1988). The recent success of this species to invade such conditions may be linked to climate warming, as it has been shown for the shrub layer in Switzerland (Walther *et al.* 2002).

Population number and size changes

Our study showed a sharp increase of the population number and size of *I. glandulifera* and *R. japonica* between 2001 and 2012 at the local scale, whereas populations of *B. davidii* only increased at some locations and *H. tuberosus* decreased overall. This overall increase is consistent with the important spread of *I. glandulifera* and *R. japonica* observed across Europe at the regional or country scale (e.g., Pysek and Prach 1993; Pysek and Hulme 2005). However, we are not aware of any study that has investigated the population number and size changes of *B. davidii* and *H. tuberosus* in Europe.

At the habitat level, high native species diversity leads to better resistance to invasive species because few resources remain available to support establishment and growth of additional species (Maskell *et al.* 2006; Eschtruth and Battles 2009). However, flood disturbances are favourable to seed germination and lead to an increase in available resources for which native and invasive species compete (Davis *et al.* 2000). Thus, invasion success will depend on the potential of species to colonise new habitats (e.g., high dispersal efficiency, high seedbank persistence, high abundance of the species in the vicinity) and to outcompete the native species (e.g., rapid growth, high faculty for regeneration; Tickner *et al.* 2001; Lake and Leishman 2004). *I. glandulifera* and *R. japonica* exhibit some of these characteristics: annual rapid growth on disturbed open soils, many persistent seeds for *I. glandulifera* (Perrins *et al.* 1993), vigorous growth and vegetative spread, with a high faculty for regeneration for *R. japonica* (Bimova *et al.* 2004). The high competitive ability (e.g., shading out native species) resulting from these characteristics could explain their population number and size increase observed between 2001 and 2012.

However, intense river flooding can also have a negative effect on the dynamics of invasive species. Kasperek (2004) noted that the cover of annual *I. glandulifera* was negatively affected by flooding occurring during the germination period (March-April) and, conversely, populations increased during dry periods. Interestingly, by comparing *I. glandulifera* densities (2001 and 2012) to hydrological patterns of the river (Fig. 2 and S7), we observed that heavy flooding occurred during spring 2001, whereas no major flooding occurred after 2007. Hence, the strong population increase may partly be related to inter-annual fluctuations of river flow. *B. davidii* may also have suffered from this decrease of flooding events, as it is known to grow rapidly on disturbed areas (Tallent-Halsell and Watt 2009) but is intolerant of shade, and natural vegetation succession may lead to its elimination in the absence of disturbance (Tallent-Halsell and Watt 2009). The low flood frequency (see Fig. S7) in recent years could explain the slower spread observed for this species between 2001 and 2012. However, given the lack of temporal data, further observations would be needed to better characterise the dynamics of this species in our study area before drawing any firm conclusion. Note that optimal conditions for *B. davidii*, characterised by stony soils, may be rarer than fine mineral sediment environments along the Venoge, thus limiting the population number of this species.

H. tuberosus grows rapidly, spreads vegetatively and has a high faculty for regeneration as well (Swanton *et al.* 1992), but showed a population size decline in our study area between 2001 and 2012. Moreover, we observed that some populations recorded in 2001 have been reduced in size and were replaced by dense populations of *I. glandulifera* or *R. japonica* in 2012. Indeed, sectors where *H. tuberosus* used to be very abundant in 2001 (i.e., sectors 26, 27, 28 and 31) correspond to the ones where *I. glandulifera* and *R. japonica* densities have increased sharply and are now very abundant (see ARF values, Fig. 5). This suggests a possible competition for suitable habitats between *I. glandulifera*,

R. japonica and *H. tuberosus*, a proposition that is additionally supported by their similar environmental requirements (Fig. 3 and S2).

Predicted distributions and derived indices

Because prevention of biological invasion at the earliest stage is more efficient than active control of well-established infestations (Leung *et al.* 2002), their spread and their effect on the native environment could be reduced by preventing the colonisation of previously preserved areas by new populations. Whereas global SDMs using topo-climatic variables at coarser resolution are useful for pre-border prevention and to depict coarse suitable areas (e.g. at the regional or national scales; Thuiller *et al.* 2005; Koop *et al.* 2012), post-borders management of biological invasions ultimately has to be performed at small spatial scales. Indeed, changes in habitat conditions can occur over short distances, especially in heterogeneous environments as alluvial areas. Thus, the very fine resolution (1 m x 1 m) used in our models allowed us to capture at the plant scale several important microhabitat variables (e.g., light availability, river sinuosity and curvature) varying significantly over short distances along rivers. Such habitat variables are better translated at such fine resolution than with a coarser resolution (e.g., 25 m or 100 m), where, for example, light availability under the canopy cannot be extracted or river sinuosity would not have any sense with a resolution exceeding river width. But these fine habitat models have to be developed only for the areas where the species is already observed (e.g., our study), or in potentially suitable areas as predicted in a first model based on topo-climatic predictors, available at coarser resolution and broader scale (e.g. temperature, precipitation).

From our predictions, we could derive PRF for the different river sectors (Fig. 5). Simultaneously, historical, recent and precise field inventories allowed us to evaluate the colonised area in the past and present, expressed as the ARF for each river sector. The combination of both indices show the colonisation potential of IPS at the local scale in the future and could be used more systematically in conservation plans along rivers to rationalise control and prioritise eradication efforts. For example, eradication could be focused on sectors presenting highly suitable conditions (i.e. high PRF) and already colonized, possibly acting as populations sources for downstream sectors (e.g. sector 32 for *I. glandulifera* and *R. japonica*, Fig. 5), or to limit the local development where very suitable conditions are present. On the other hand, observation efforts could be prioritize in suitable (i.e. high PRF) but unfilled sectors (i.e low ARF), such as sector 16 for *I. glandulifera* or sector 35 for *R. japonica* (Fig. 5).

In our field study area, suitable habitats are still largely available for each IPS, supporting the prediction that the observed increase in invasive species will continue in coming years. These predicted areas represent minimum scenarios, as biological invasions are ongoing phenomena and thus it is impossible to postulate if the species have reached their environmental equilibrium (i.e. they have colonized all possible suitable environmental conditions). Although the massive increase of populations for some species between 2001

and 2012 suggests they present a more exhaustive niche along the Venoge River, only future inventories will allow testing if species have reached their niche equilibrium. In the meanwhile, our approach provides a spatially explicit information about current stage and future risk of invasions that should help the management of IPS at a very local scale.

Acknowledgments

Thanks to Stuart Lane and Nathalie Diaz for their assistance with soil sample analysis, Anne Freitag and the Zoological Museum of Lausanne for the loan of the spherical densiometer, and Robin Gray and Jérôme Pellet for the helpful comments.

REFERENCES

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

Araújo, M.B., New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution,* **22,** 42-47.

Barney, J.N., Tharayi,l N., DiTommaso, A., Bhowmik, P.C. (2006) The biology of invasive alien plants in Canada *Polygonum cuspidatum* Sieb. & Zucc. [= *Fallopia japonica* (Houtt.) Ronse Decr.]. *Canadian Journal of Plant Science,* **86,** 887-905.

Beerling, D.J., Bailey, J.P., Conolly, A.P. (1994) *Fallopia japonica* (Houtt.) Ronse Decraene *Journal of Ecology,* **82,** 959-979.

Beerling, D.J., Perrins, J.M. (1993) *Impatiens glandulifera* Royle (*Impatiens* Roylei Walp.), Journal of Ecology, **81,** 367-382.

Bimova, K., Mandak, B., Kasparova, I. (2004) How does *Reynoutria* invasion fit the various theories of invasibility? *Journal of Vegetation Science*, **15,** 495-504.

Davis, M.A., Grime, J.P., Thompson, K. (2000) Fluctuating resources in plant communities: a general theory of invasibility. *Journal of Ecology*, **88,** 528-534.

Dray, S., Dufour, A.B. (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statitistical Software,* **22,** 1-20.

Ehrenfeld, J.G. (2003) Effects of exotic plant invasions on soil nutrient cycling processes. *Ecosystems,* **6,** 503-523.

Elith, J., Leathwick, J.R., Hastie, T. (2008) A working guide to boosted regression trees. *Journal of Animal Ecology,* **77,** 802-813.

Elith, J., Phillips, S.J., Hastie, T., Dudik, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions,* **17,** 43-57.

Eschtruth, A.K., Battles, J.J. (2009) Assessing the relative importance of disturbance, herbivory, diversity, and propagule pressure in exotic plant invasion. *Ecological Monograph,* **79,** 265-280.

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.

Friedman, J., Hastie, T., Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting. *Annals of Statistics,* **28,** 337-407.

Glutz Von Blotzheim, U.N. (1988) *Handbuch der Vögel Mitteleuropas. Band 11/II.* AULA-Verlag, Wiesbaden.

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

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Ferrier, S., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters,* **16,** 1424-1435.

Heinänen, S., Erola, J., Von Numers, M. (2012) High resolution species distribution models of two nesting water bird species: a study of transferability and predictive performance. Landscape Ecology. **27,** 545-555.

Hijmans, R.J., Phillips, S., Leathwick, J., Elith J. (2011) Dismo: Species distribution modeling. R package version 0.7- 17. http://cran.r-project.org/package=dismo.

Hill, M.O., Smith, A.J.E. (1976) Principal component analysis of taxonomic data with multi-state discrete characters. *Taxonomy,* **25,** 249-255.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling,* **199,** 142-152.

Hulme, P.E. (2003) Biological invasions: winning the science battles but losing the conservation war? *Oryx,* **37,** 178-193.

Kasperek, G. (2004) Fluctuations in numbers of neophytes, especially Impatiens glandulifera, in permanent plots in a west German floodplain during 13 years. *NEOBIOTA,* **3,** 27-37.

Keller, C.A., Hill, M., Vollmer, M.K., Henne, S., Brunner, D., Reimann, S., O'Doherty, S., Arduini, J., Maione, M. & Ferenczi, Z. (2011) European emissions of halogenated greenhouse gases inferred from atmospheric measurements. *Environmental science & technology,* **46,** 217-225.

Koop, A.L., Fowler, L., Newton, L.P., Caton, B.P. (2012) Development and validation of a weed screening tool for the United States, *Biological Invasions,* **14,** 273-294.

Lachat, T., Pauli, D., Gonseth, Y., Gregor, K., Scheidegger, C., Vittoz, P., Walter, T. (2010) *Wandel der Biodiversität in der Schweiz seit 1900. Ist die Talsohle erreicht?* Haupt Verlag. Berne.

Lake, J.C., Leishman, M.R. (2004) Invasion success of exotic plants in natural ecosystems: the role of disturbance, plant attributes and freedom from herbivores. *Biological Conservation,* **117,** 215-226.

Landolt, E. (2010) *Flora indicativa: Ecological indicator values and Biological attributes of flora of Switzerland and the Alps.* Haupt, Bern.

Leung, B., Lodge, D.M., Finnoff, D., Shogren, J.F., Lewis, M.A. & Lamberti, G. (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society of London. Series B: Biological Sciences,* **269,** 2407-2413.

Maskell, L., Bullock, J., Smart, S., Thompson, K. & Hulme, P. (2006) The distribution and habitat associations of non-native plant species in urban riparian habitats. *Journal of Vegetation Science,* **17,** 499-508.

McCullagh, P., Nelder, J.A. (1989) *Generalized linear models, 2nd ed.* Chapman and Hall, London.

Miller, N.P., Matlack, G.R. (2010) Population expansion in an invasive grass, Microstegium vimineum: a test of the channelled diffusion model. *Diversity and Distribution,* **16,** 816-826.

Perrins, J., Fitter, A., Williamson, M. (1993) Population biology and rates of invasion of three introduced Impatiens species in the British Isles. *Journal of Biogeography,* **20,** 33-44.

Phillips, S.J., Anderson, R.P., Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions*. Ecological Modelling,* **190,** 231-259.

Pysek, P., Chytry, M., Pergl, J., Sadlo, J. & Wild, J. (2012) Plant invasions in the Czech Republic: current state, introduction dynamics, invasive species and invaded habitats. *Preslia,* **84,** 575-629.

Pysek, P., Hulme, P.E. (2005) Spatio-temporal dynamics of plant invasions: Linking pattern to process, *Ecoscience,* **12,** 302-315.

Pysek, P., Prach, K. (1993) Plant invasions and the role of riparian habitats: a comparison of four species alien to central Europe. *Journal of Biogeography,* **20,** 413-420.

Raunkiær, C. (1934) *The life forms of plants and statistical plant geography*. Oxford University Press, Oxford.

Renöfält, B.M., Jansson, R., Nilsson, C. (2005) Spatial patterns of plant invasiveness in a riparian corridor. *Landscape Ecology,* **20,** 165-176.

Richardson, D.M., Allsopp, N., D'Antonio, C.M., Milton, S.J. & Rejmanek, M. (2000) Plant invasions - the role of mutualisms. *Biological Reviews,* **75,** 65-93.

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

Swanton, C.J., Cavers, P.B., Clements, D.R., Moore, M.J. (1992) The biology of Canadian weeds, *Helianthus tuberosus* L. *Canadian Journal of Plant Science,* **72,** 1367-1382.

Tallent-Halsell, N.G., Watt, M.S. (2009) The invasive *Buddleja davidii* (Butterfly bush), *Botanical Reviews,* **75,** 292-325.

Thuiller, W., Lafourcade, B., Araujo, M. (2010) *Presentation manual for BIOMOD*. University Joseph Fourier, Laboratoire d'écologie Alpine, Grenoble.

Thuiller, W., Richardson, D., Pysek, P., Midgley, G., Hughes, G. & Rouget, M. (2005) Niche-based modelling as a tool for predicting the global risk of alien plant invasions. *Global Change Biology,* **11**. 2234-2250

Tickner, D.P., Angold, P.G., Gurnell, A.M., Mountford, J.O. (2001) Riparian plant invasions: hydrogeomorphological control and ecological impacts. Progress in Physical Geography, **25,** 22-52.

Václavík, T., Meentemeyer, R.K. (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? Ecological Modelling, **220,**3248-3258.

Venables, W.N., Ripley, B.D. (2002) *Modern applied statistics with S. Fourth edition*. Springer, New York.

Vicente, J., Randin, C.F., Goncalves, J., Metzger, M.J., Lomba, A., Honrado, J. & Guisan, A. (2011) Where will conflicts between alien and rare species occur after climate and land-use change? A test with a novel combined modelling approach. *Biological Invasions,* **13,** 1209-1227.

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

Ward, J.V., Tockner, K., Arscott, D.B., Claret, C. (2002) Riverine landscape diversity, *Freshwater Biology,* **47,** 517-539.

Wyse, D.L., Young F.L., Jones, R.J. (1986) Influence of Jerusalem Artichoke (*Helianthus tuberosus*) density and duration of interference on Soybean (Glycine max) growth and yield. *Weed Science,* **34,** 243-247.

SUPPORTING INFORMATION

Appendix S1 Methodological details about index calculations

Figure S1 Evolution of the size of the IPS populations between 2001 and 2012

Figure S2 Mixed multivariate analysis (MMA) showing the environmental distance between the IPS niche centroids (2 dimensions)

Figure S3 Response curves for the five most influential predictors of the species occurrence models

Figure S4 Response curves for the five most influential predictors of the species abundance models

Figure S5 Mapped scores of habitat suitability for the occurrence and abundance model of *I. glandulifera* along a portion of the Venoge River.

Table S1 Environmental distance between the IPS niche centroids

APPENDIX S1: METHODOLOGICAL DETAILS ABOUT INDEX CALCULATIONS

Soil texture

Soil texture was estimated visually (in percentage) at a depth of 5 cm for the following four classes: organic matter, fine mineral sediments (< 2 mm), small stones (2-75 mm) and large stones (> 75 mm). These visual estimations were then corrected using a standardised reference. The latter was obtained from 10 representative soils observed under our populations that were sieved and weighed. Organic content in the fine mineral sediment fraction (< 2 mm) was measured following Allen (1974). When compared to visual estimations, these measurements showed a constant overestimation of the visual organic fraction combined to an overall underestimation of the sediment fraction. However, a good correlation between the visual estimation of organic fraction and laboratory measurements (linear regression: $R^2 = 0.83$; p-value < 0.001) allowed us to correct the estimated organic and sediments values.

Distances

Distances to river, roads and railway were extracted by using the Euclidean distance tool in ArcGIS and densities of industries and habitations were calculated on a circular window of 500 m radius by using land cover data obtained from the Swiss Federal Office of Statistics (http://www.geostat.admin.ch) for the year 2009.

River sinuosity and curvature

River sinuosity was calculated for each meter along the river as the ratio between the 100 m separating two points along the watercourse and the shortest straight path connecting them. Curvature was measured as the difference between Pi (π) and the angle (in radian) formed by two successive segments connecting one point 20 m before and one point 20 m after the focal points. Focal points were distributed each meter along the river flow on both right and left river sides. These two river indices were extended to the whole study area by using the Euclidean allocation tool in ArcGIS.

Canopy density

Many techniques were developed to assess estimates of the forest canopy density. Jennings *et al.* (1999) have distinguished two types of measurements of the canopy density: the canopy closure defined as the proportion of the sky hemisphere covered by high shrubs and trees when viewed from a single point, and the canopy cover defined as the proportion of the soil surface covered by the vertical projection of the tree layer. In our study, we measured the canopy density with two techniques. First, the field canopy closure (CC_{Field}) was estimated with a convex spherical densiometer (Lemmon 1956) at each sample sites (populations and randomly selected sites). The remotely sensed canopy

cover (CC_{LiDAR}) was derived from the LiDAR DEM acquired between 2001 and 2002 at 1 m resolution (MNT-MO/MNS; OIT). CC_{LiDAR} was calculated following Hopkinson and Chasmer (2009), with a focal circular window as the ratio between non-ground returns and total returns. Since we used CC_{LiDAR} as a proxy of CC_{Field} for the spatial projection, we fixed the circular window at 8 m radius because it provided the best linear relationship between the two metrics (linear regression: $R^2 = 0.42$).

REFERENCES

Allen, S.E. (1974) *Chemical analysis of ecological materials*. Blackwell Science, Oxford.

Hopkinson C., Chasmer L. (2009) Testing LiDAR models of fractional cover across multiple forest ecozones, *Remote Sensing of Environment,* **113,** 275-288.

Jennings S.B., Brown N.D., Sheil D. (1999) Assessing forest canopies and understorey illumination, canopy closure, canopy cover and other measures, *Forestry,* **72,** 59-73.

Lemmon P.E. (1956) A spherical densiometer for estimating forest overstory density, *Forestry Sciences,* **2,** 314-320.

Table S1. Environmental distance between the niche centroids of the species and the other species centroids or the random points (Background) calculated on the three first axes of a mixed multivariate analysis (MMA) built with 17 environmental predictors and explaining 37.06 % of the data variation (Fig. 3 and see Fig. S2).

	Background	*B. davidii*	*H. tuberosus*	*I. glandulifera*	*P. laurocerasus*
B. davidii	1.70	-	-	-	-
H. tuberosus	1.82	1.01	-	-	-
I. glandulifera	1.33	1.18	0.65	-	-
P. laurocerasus	3.26	2.60	3.30	3.15	-
R. japonica	1.36	1.00	0.83	0.42	2.72

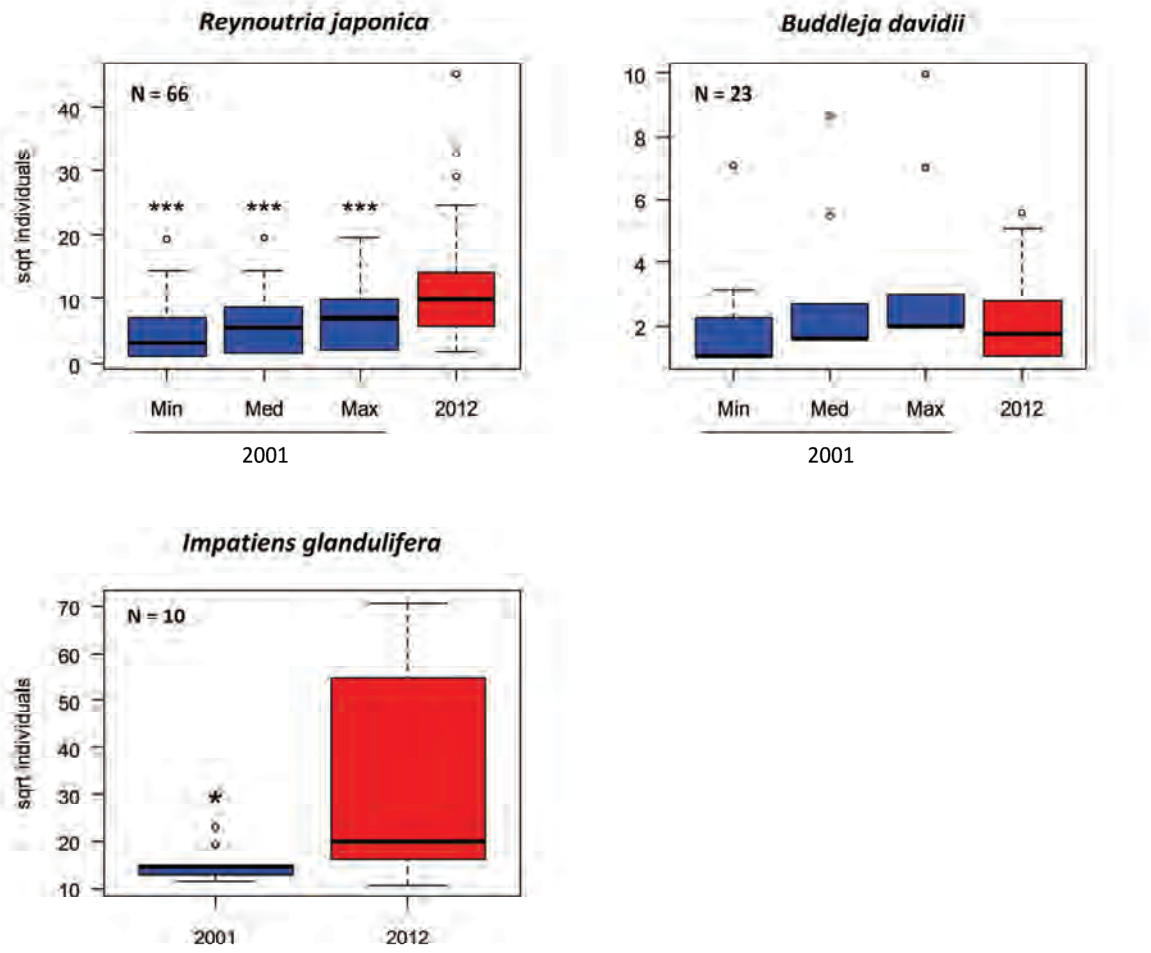


Figure S1. Evolution of the size (number of individuals) of the IPS populations between 2001 and 2012. R. japonica and B. davidii population size for 2001 were extracted from the sizes classes (classes 1 to 5) following three conversions: minimum (Min), median (Med) and maximum (Max) number of individuals in size classes. One outlier of R. japonica in 2012 (i.e. 2020) was not represented. I. glandulifera and R. japonica are the species presenting a significant increase of their population size between 2001 and 2012 (*, $p < 0.05$; ***, $p < 0.001$). The intermediate line is the median and the boxes are limited by 1st and 3rd quartiles. Note that the scales are different between species.

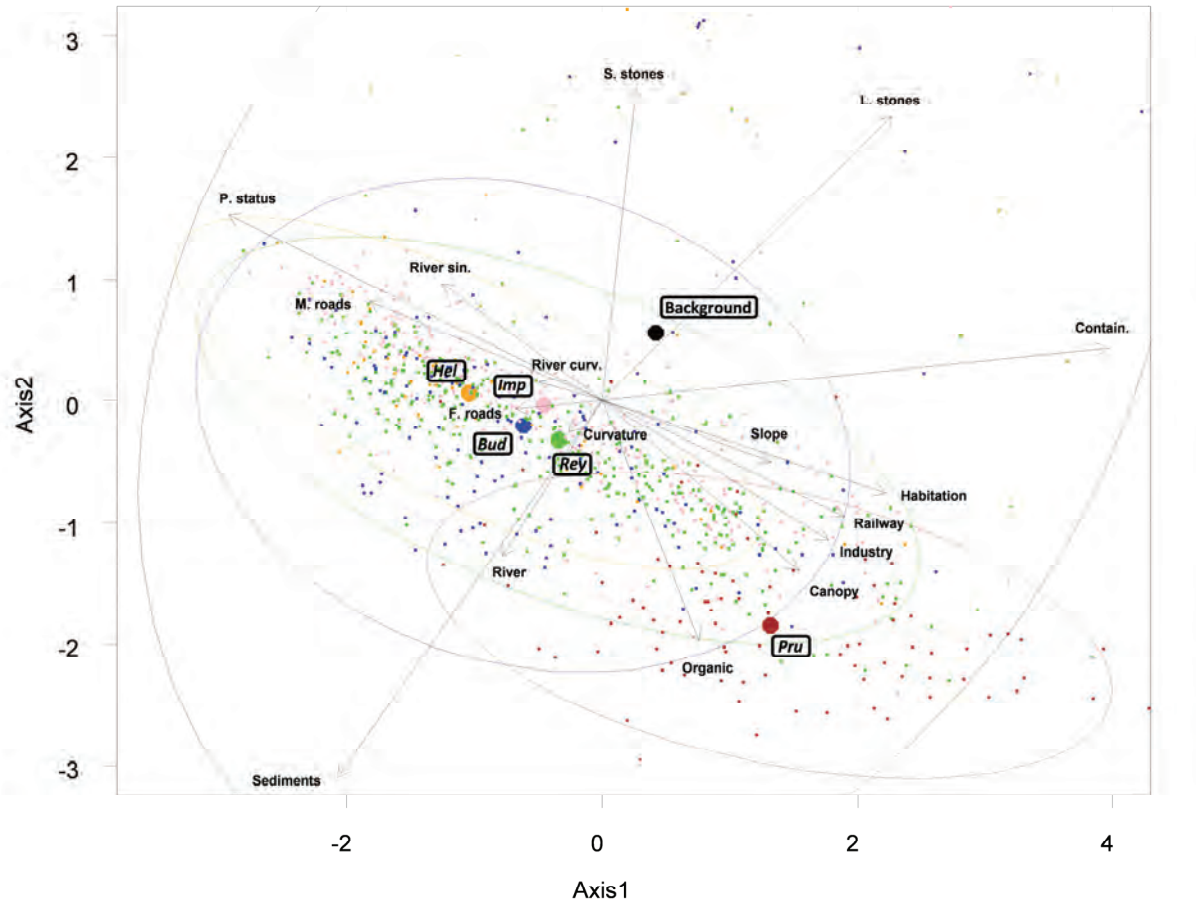


Figure S2. Mixed multivariate analysis (MMA) built with 17 ecological predictors and showing the environmental distance between the niche of *R. japonica* (Rey), *I. glandulifera* (Imp), *B. davidii* (Bud), *H. tuberosus* (Hel), *P. laurocerasus* (Pru) and the random points (Background). The two first axes explained 27.86 % of the data variation (axis 1: 15.09 %; axis 2: 12.77 %). The points correspond to the data and the length of the vectors represents the magnitude of the correlation between the variables and the axes. Variable abbreviations are given in Table 2. The niche of *P. laurocerasus* is well separated from the other species.

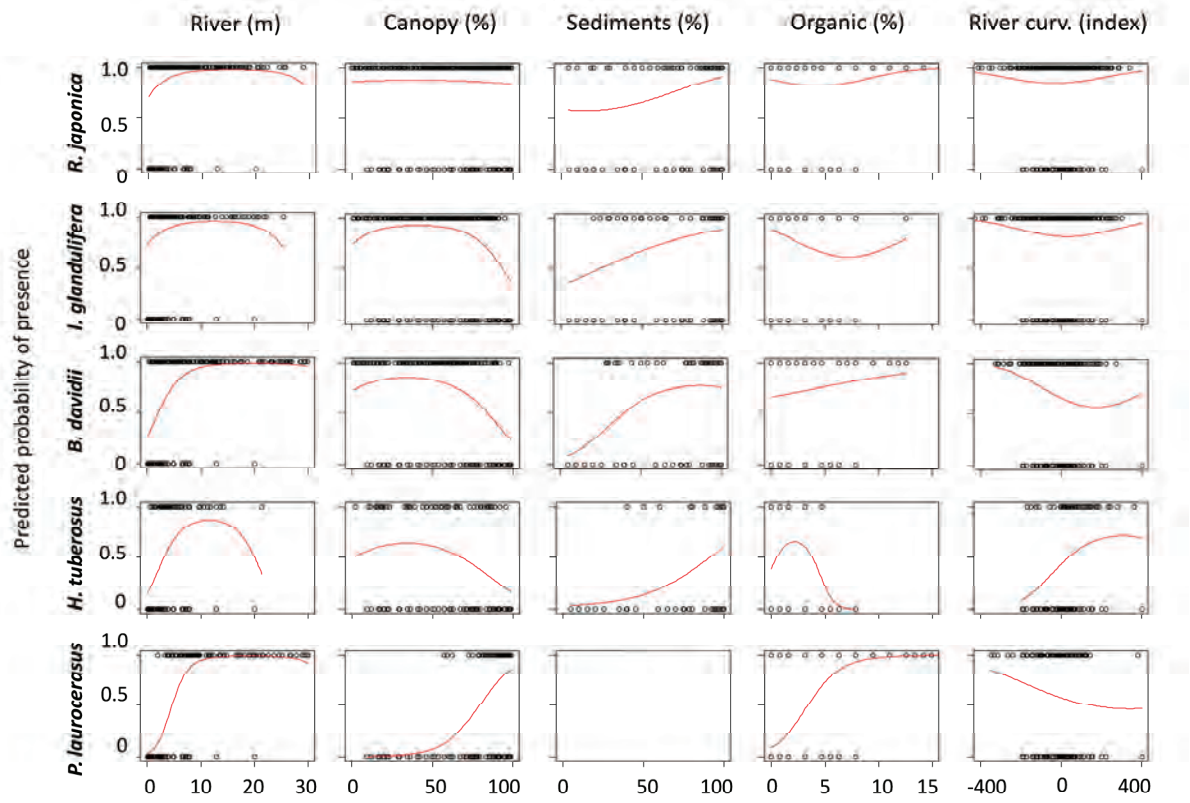


Figure S3. Response curves for the five most influential predictors of the species occurrence models. The curves demonstrate the influence of the variable on the GLM species-occurrence prediction probability. The points represent the variables values for the absence data (bottom of the graph) and the species observations (top of the graph). The sediments variable was removed from the *P. laurocerasus* model to avoid data overfitting. GBM and MAXENT response curves presented similar patterns of influence (data not shown).

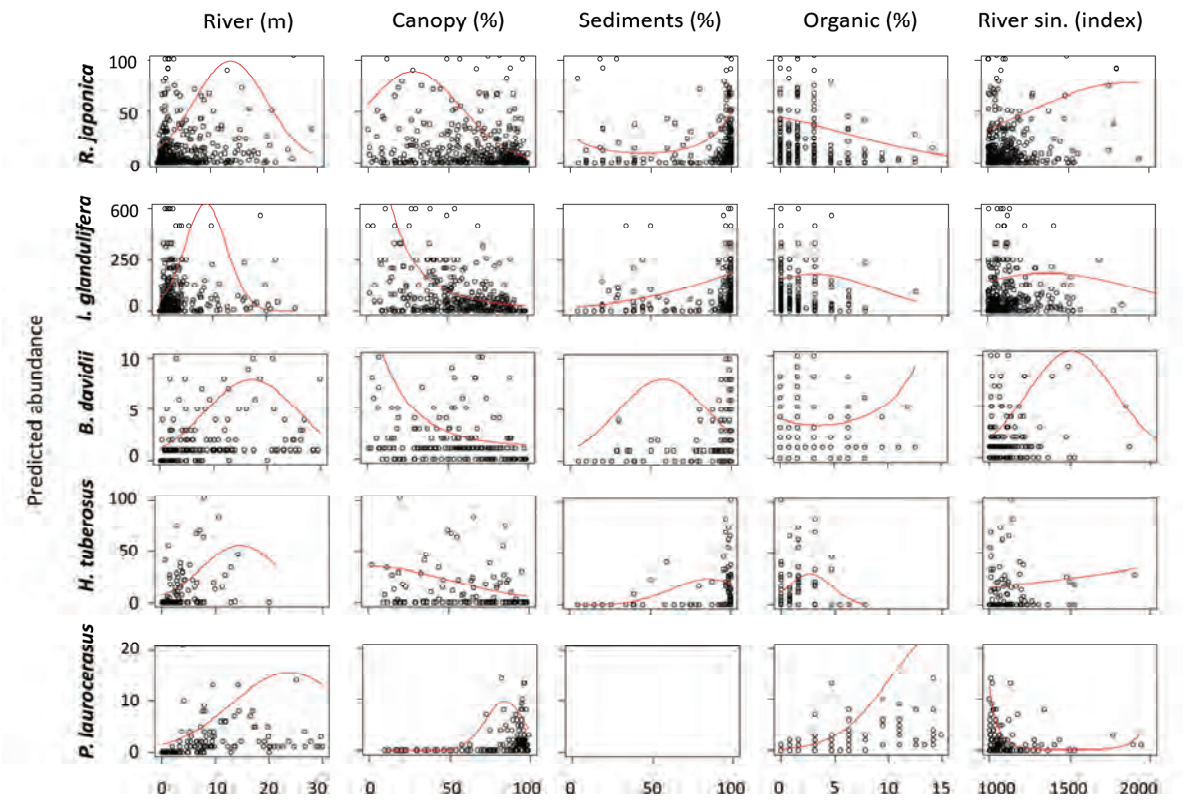


Figure S4. Response curves for the five most influential predictors of the species abundance models. The curves demonstrate the influence of the variable on the GLM species-abundance prediction. The points correspond to the observations. The sediments variable was removed from the *P. laurocerasus* model to avoid data overfitting. GBM response curves presented similar patterns of influence (data not shown). Note that the scales on y-axis (number of individuals per population) are different between species.

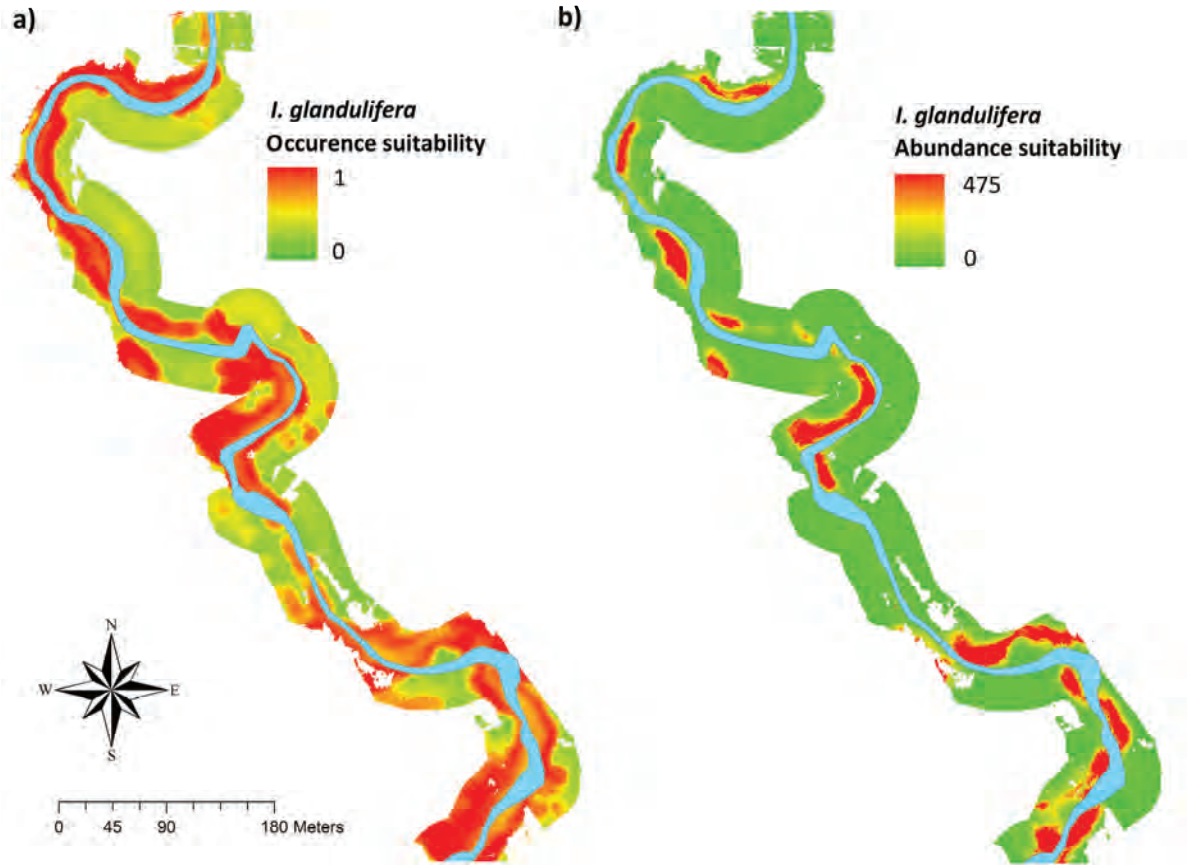


Figure S5. Mapped scores of habitat suitability for the (a) occurrence and (b) abundance model of *I. glandulifera* along a portion of the Venoge River (sectors 27 to 29). Scores were obtained by averaging each single model projections (ensemble function). High scores indicate that the habitat is favourable for the presence or a high abundance of the species (occurrence and abundance model, respectively). The distribution pattern of the suitable habitats indicates that the species occurrence model is mainly influenced by the light availability and that the species abundance model is mainly influenced by the light availability and the river curvature (Fig. 4).

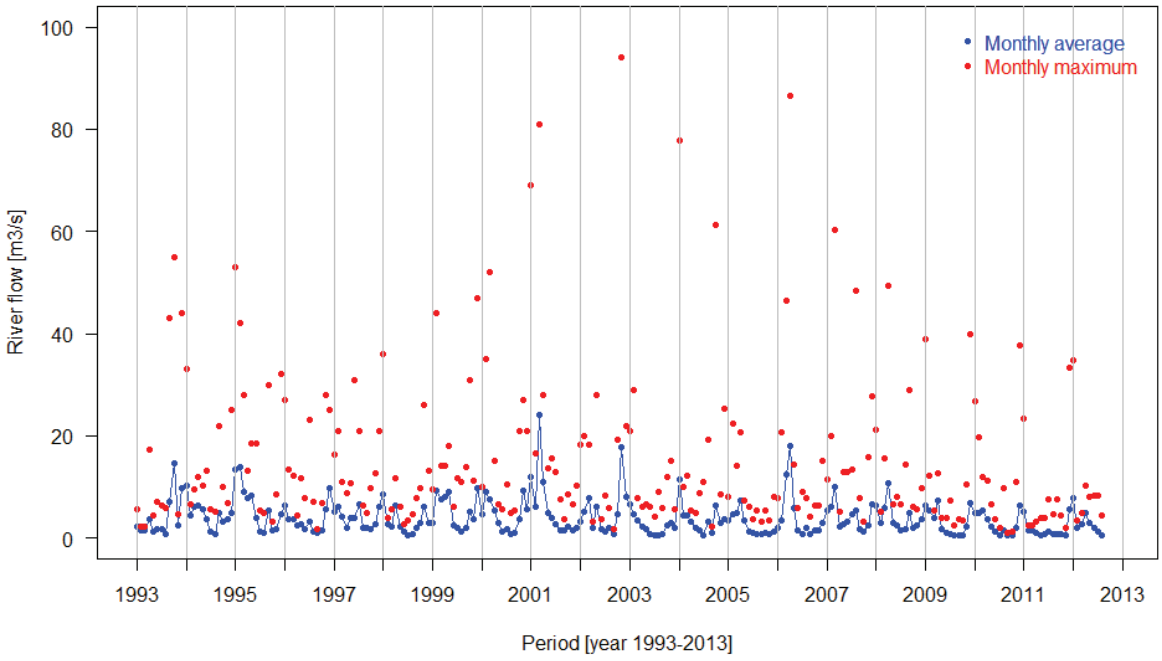


Figure S6. Monthly average (connected blue dots) and maximum (red dots) of the Venoge River flow (m3/s) for the period 1993-2013. Data collected by a hydrological station (Venoge - Ecublens, Les Bois 2432) of the Swiss Federal Office for the Environment (http://www.hydrodaten.admin.ch).

SYNTHESIS & MOVING FORWARD

Biological invasions provide one of the rare opportunities to compare distributions of the same species in two (or more) distinct study areas, thus analogous to a worldwide experiment (Sax *et al.*, 2007). My thesis essentially used invasive species to investigate the conservatism of their realized niche through space, and assess our related capacity to anticipate biological invasions through the use of niche models. The implications of the results presented in my thesis thus reach a broader problematic than biological invasions only, expanding to our ability to understand and predict species response to rapid global change in general. The three main axes of my thesis were 1) to provide a framework allowing a proper quantification and interpretation of niche changes, 2) testing for niche conservatism at the general level with the aim of overstepping contradicting case studies, and 3) assessing impacts of the ongoing global change on the response of species distribution for a better anticipation and thus better conservation strategies. In the following sections I discuss the advances but also the limitations of my results and possible alternative to move forward.

DISENTANGLING NICHE CHANGES CATEGORIES

My thesis contributed to implement a quantitative framework particularly adapted to the problem of species with ongoing colonization, such as biological invaders. My approach was based on the work of Broennimann *et al.*, 2012 (Appendix 1), which itself was deeply inspired by the work of Warren *et al.*, 2008 and its aim to properly define and quantify niche conservatism. The advances proposed by Broennimann *et al.*, 2012 were twofold. First, by translating Warren's concepts to the environmental space depicted by ordination techniques and thus taking into account the differences in environmental availability between the two study areas, tests and metrics were less subject to biases caused by the geographical components. Second, it used virtual species to assess that a PCA calibrated on the whole extent areas was the most accurate way to quantify and test niche changes, and provided better views on which analytical tool (ordination or reciprocal ENM) should be used for assessing niche changes. However, it still remained based on an overall niche change measurement (i.e. niche overlap), without disentangling the different types of niche changes that can be observed during an invasion (Chapter 1.1 and 1.2).

Implementing the unfilling/expansion categorization provided insightful perspectives: for example, I showed that most niche changes were due to unfilling rather than expansion in Holarctic plant invaders (Chapter 1.2). I focused mostly on the expansion index because this latter was easier to interpret with an ENM approach only. Interpreting unfilling in the case of biological invasions is hazardous as most invasive species show ongoing dispersal (Hulme, 2009). Although it is possible that biological phenomena explain invasive species' absence from suitable area in its invaded range (e.g. non-preadapted ecotypes, competitions with other native or exotic species), there is a higher probability that unfilling reflects species dispersal limitation. On the other hand, if species are assumed to be at environmental equilibrium in their native range and expand towards environments that are

available but not colonized in the native range, it is a strong support for competitors release, biotic facilitation or evolution of increased competitive ability. Of course, the assumption that species is at its equilibrium in the native range has to be considered with care. Nevertheless, unfilling remains a promising dimension to explore but requires population genetics, experiments or implementation of dispersal models (Appendix 2 and 3) which are costly to apply to numerous species at the same time, as I did for expansion.

NICHE CHANGES CATEGORIES AND EVALUATION OF SDM

It is also important to consider this asymmetry between these two components, unfilling and expansion, when predicting invasive species assumed to be still expanding (i.e. ongoing dispersal). Commonly used SDM's evaluators like AUC or TSS focus on both the rates of presences correctly predicted (sensitivity) and the rate of absences correctly predicted (specificity). It was thus delicate to evaluate, with these latter criteria, a species known not to have colonized all its suitable areas. In my thesis, I implemented the evaluation through the Boyce index and the sensitivity, two predictors based essentially on presences, which thus proved particularly relevant in the invaded range. The Boyce index (B; Hirzel *et al.*, 2006) was popularized through Biomapper (Hirzel *et al.*, 2007), an SDM software contrasting presences from the remaining environment of the study area, but was not commonly used outside this tool so far. B shares many similarities with AUC. It is analogous to a correlation, can be threshold independent, strongly depends on the study area and may be difficult to interpret (Hirzel *et al.*, 2006, Phillips and Elith, 2010). This is why I always combined it with sensitivity, which is easier to interpret, analogous to expansion (sensitivity would correspond to Se = 1 - E in the geographical space) but requires a threshold to binarize data. The choice of the threshold is always arbitrary and depends on the questions behind the use of SDMs. Thus by using presence only evaluation criteria, I adapted the way to evaluate SDM specifically for invasive spreading species. The field of simple and easily interpretable presence-only evaluators is still poor. Recently, the POC-plot was proposed as a promising and unbiased way to evaluate presences only (Phillips and Elith, 2010). It represents indeed the same graphical context within which the Boyce index is calculated. Without a quantitative measure associated, it requires subjective interpretations and may be difficult to implement for large species dataset as I used in my thesis, thus using B instead. A possible way to follow up would be to combine several evaluators, all providing different information, in a consensual way as I attempted in Chapter 1.3.

DISENTANGLING CLIMATE AVAILABILITY

An important dimension raised with my approach to quantify niche changes is the consideration of differences in climate availability between the two study areas (as inspired from Broennimann *et al.* 2012; Chapter 1.1). How should we consider environments that

are not available in the native range but present in the invaded range? This problematic also occur when SDMs are applied with climatic scenarios containing not yet existing climates (Fitzpatrick and Hargrove, 2009). Popularized by the implementation of the multivariate environmental similarity surface (MESS) in the software MAXENT (Elith *et al.*, 2010) and pointed out by several studies underlining the importance of gradients truncations on SDM outputs and interpretation (e.g. Thuiller *et al.*, 2004, Broennimann *et al.*, 2012), this problematic has only emerged recently. In my thesis, I mostly focused on environments that were available and comparable between the two ranges. I did not aim at projecting species niche outside the range of conditions available in their native range. Yet, exploring the niche of invasive species in non-analog invaded areas should provide interesting complementary perspectives. First, it can provide valorous insights about the nature of dispersal limitation in the native range (intrinsic or extrinsic; Baselga *et al.*, 2012). Second, it can inform about the physiological relevance of SDMs and thus the predictive power of simple models correlating occurrences to environmental factors. This latter point can only be done through the coupling of SDMs with mechanistic physiological models (e.g. Kearney and Porter, 2009).

DELIMITING THE STUDY AREA

The differences in climate availability strongly depend on the study area delimitations, also called extent or background area. The extent does not only impacts available climates but also SDM predictions (Barve *et al.*, 2011) and some tests of niche conservatism (e.g. Rödder and Engler, 2011). In my thesis, I delimited study areas as the biogeographical units (realms or ecoregions) covered by the species distribution. This approach seemed relevant when working with (quasi) exhaustive species distribution datasets at relatively coarse resolution. I intended to test the effect of different ways to delimit the study area, using the Holarctic plant invaders dataset. In the same way as I did for variable selection (Chapter 1.3), invaded ranges represent a valuable opportunity to test the transferability of SDMs. However, the coarse resolution of the distribution data (50 km) prevented to include finer levels such as ecoregions or land-cover, and thus would not be very informative. Using finer exhaustive resolution data about native and invaded species ranges could provide valuable information about the optimal way to delimit the study areas. I started to investigate the effect of the calibration area in Appendix 3.

Note that all these methodological questions, from SDM evaluation to impact of the delimitation of the study areas, through the extrapolation of SDM to non-analog climates, can be complemented by a “virtual ecologist” approach (Zurell *et al.*, 2010). Simulating species distributions and assessing the predictive power of various statistical methods, biases due to spatial autocorrelation, sampling design or due to the number of samples included in the study, was possible through this approach (Broennimann *et al.*, 2012, Appendix 4).

PREDICTING INVASIVE SPECIES DISTRIBUTION AT DIFFERENT SCALES

In an analog way as species traits diversity (Whitaker 1975), the scale of the species niche can be disentangled into two components: i) the α niche, corresponding to a small-scale components of the niche that differ among co-occurring species, typically at the micro-habitat level, whereas ii) the β niche is defined by macrohabitat and climate factors related to larger scale distributions (Silvertown 2006). Although I worked at these two levels in my thesis, I tested niche conservatism between native and invaded range only at the β level. At such scale, the moderate (Chapter 1.4) to strong (Chapter 1.2 and 1.3) signal of realized niche conservatism between species' native and invaded ranges supports the overall use of SDMs to predict species distribution under global changes, although tests of niche changes should be used to detect species for which transferability is likely to be limited. It would evidently be desirable to build more mechanistic models based on species' physiology, phenology, traits, dispersal rate and population genetics, but the latter factors are more difficult to implement for a wide range of species, and thus their use is more limited if general conclusions have to be drawn (See Appendix 2 for a case studies and Appendix 5). In Chapter 1.3, I showed that it is possible to predict the invaded ranges of 96% of the most widespread Holarctic plant invaders by using the most predictive variables dataset. In mammals (Chapter 1.4), the signal, although significant, was weaker possibly due to their endothermy, making their realized niche the results of dispersal and competition more than physiological tolerances (Dormann *et al.*, 2010). The fact that widespread species are predictable through space using SDMs supports the use of these models to also predict their niche through time, under various climate change scenarios. Indeed, invasive species and global warming are jointly affecting species distributions and biodiversity and thus generating major changes in biotic interactions. In this sense, my thesis combined these two problematics in Chapter 2.1 and showed that the predicted responses of plant invaders to climate change are roughly the same as those predicted for the alpine flora: climbing up the mountains.

Although such coarse resolution is typically used for pre-border weed risk assessments (Koop *et al.*, 2012), it may not be pragmatically useful for conservation managers applying conservation planning at a local scale, closer to the α niche. Testing SDM transferability at fine scale with the same approach as I did at coarse scale was not possible for 2 reasons. First, it required to cover the entire species niche, but with finer occurrences and environmental datasets. Climatic datasets have become now accessible at a relatively fine scale (e.g. worldclim at 1 km, Hijmans *et al.*, 2005) whereas global and exhaustive occurrence datasets are far more difficult to obtain. Atlas data have a coarse resolution and global database such as GBIF are biased towards Western Europe and North America (Boakes *et al.*, 2010) and includes exponentially less occurrences at 1 km scale and below. Second, factors affecting the α niche (e. g. land-cover, light availability, topography, distances to roads, rivers or cities) often exists at the national or continental scale (e. g. CORINE landuse categories for European Union) but requires a substantial effort to be compiled and standardized at global scale.

As the α niche is hierarchically nested into the β niche, their modeling can be also hierarchical. A hierarchical approach integrating information about the both levels with different types of predictors may be a safer way than simply downscaling macroclimatic SDMs at a finer regional scale (McPherson *et al.*, 2006). The assumption that the same rules relating coarse species distribution to coarse environmental data also apply to finer distribution and environmental data may not hold (Menke *et al.*, 2009). Typically, the influence of biotic interactions on the modeled niche can vary with scales and resolutions. It is assumed that the coarser the resolution of ENM, the less important are the biotic interactions in the shape of the realized niche (Prinzing *et al.*, 2002, Kissling *et al.*, 2012, see the Eltonian Noise Hypothesis in Soberón & Nakamura, 2009). It is possible to include biotic interactions in ENM (see Kissling *et al.*, 2012 for a review) and such inclusions have shown tremendous increase of predictive power, although their effect tend to decrease with coarser resolution (e.g. from a 10 to a 40 km resolution in Heikkinen *et al.*, 2007). It is then possible that the high rate of niche conservatism found at coarse resolution may be lower at finer scale where the biotic component of the realized niche becomes more important. However, without exhaustive fine resolution data, it is difficult to assess niche conservatism at finer scale. The approach I chose in the Chapter 2.1 of my thesis to predict species at finer scale was to compare and compile coarse resolution SDMs calibrated at worldwide scale with high resolution SDMs calibrated at local scale, using partially different environmental predictors at each step. Additionally, information from global SDMs was used to select pseudo-absences in the local SDMs (Gallien *et al.*, 2012). Coarse and fine SDMs could be compiled in a more complex way, using a new hierarchical Bayesian framework to integrate the information of the β niche to the modeling of the α niche as a prior information. Bayesian framework showed promising results in downscaling (Keil *et al.*, 2013) but the computational efforts make this approach still reserved for smaller study areas than the one used in this thesis. Nevertheless, the Chapter 3.1 showed promising results for very fine scale species predictions, based only on topographic, disturbances and light-availability factors. To my knowledge, no SDM was previously applied at such a fine scale in regards to the covered area. Testing their transferability on similar study areas would be the next step to conclude about their predictive ability at finer resolution.

IS THE ENVIRONMENTAL NICHE ONE MORE TRAIT?

Niche and traits analyses are often seen as separated and alternative approaches to study invasive species, even though claims exist that these two approaches should be combined (Moles *et al.*, 2008, Hulme and Barrett, 2013). Although I did not combine traits analyses and niche models in a proper defined framework, I attempted to integrate the traits components as much as possible in my ecological niche analyses. When working with numerous species, as I did, it becomes impossible to describe individual responses for each species. Thus, seeking for trends among traits allowed for a more biological perspective on the observed niche responses such as: “who are the mountain climbers?”

(Chapter 2.1), “who are the predictable species” (Chapter 1.3). Further exciting questions are raised by looking for the causes of niche shifts. Relating species traits to species niche shift may open promising perspectives to understand the causes of niche shifts/stasis (Strubbe *et al.*, 2013, Appendix 5). For example, I found the same negative relationship between native niche size and species ability to shift (Appendix 5) as we found for Holarctic invaders (Fig. 1).

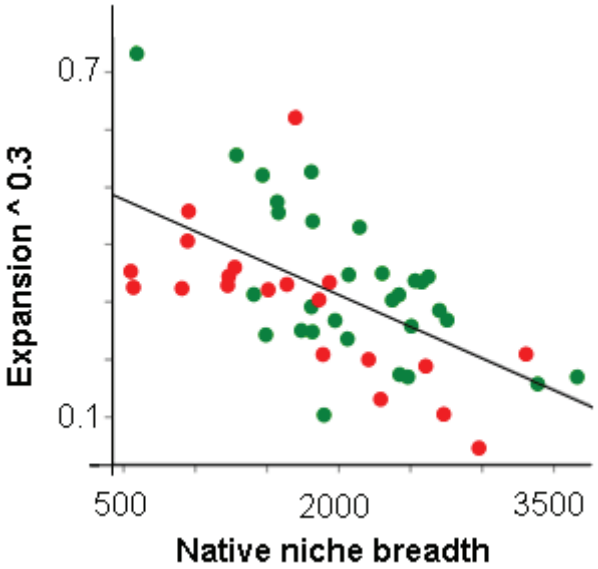


Figure 1. Relationship between native niche breadth and expansion. Red and green dots represent American and Eurasian species respectively. $R^2 = 0.3$

When ENM is not sufficient to predict correctly success of introduction in mammals (Chapter 1.4), native niche could considered as one among numerous other traits related to species’ morphology. Combining SDM and traits analyses for numerous species allows a better understanding of the biological processes driving species distributions.

CONCLUSIONS

Although affected by several draw-backs, the ENM approach has one huge advantage: it allows sampling hundred of species at the worldwide scale. My thesis used this strength to overstep case studies and learn from generalities (Chapters 1.2, 1.3, 1.4, 2.1), which would have been impossible with mechanistic or dispersal models. By formulating a proper framework to compare realized niche between native and invaded ranges for numerous species (Chapters 1.2 and 1.4), I showed that climate is a significant factors for predicting biological invasions (Chapters 1.3, 1.4). More generally, this supports the continued use of ENM approaches to predict how species distribution may be affected by the rapid global change that Earth is currently experiencing (Chapter 2.1 and 2.2). It also illustrates that the naturalizations of exotic species is only one of the numerous component of this human

induced global change (Chapter 2.1 and 2.2). By combining different approaches at different scales, my thesis provides the tools for a better understanding of the components behind invasive species distribution, allowing better anticipation and thus conservation policies to lower the impact of the Anthropocene on biodiversity (Chapter 2.1, 2.2 and 3.1). It also paved the way for to address remaining challenges: modeling at global scale to predict a fine local resolution, understanding the niche in non-analog climate, converging to more mechanistic and physiological of the species studied to bridge the species trait's with species' niche.

REFERENCES

Austin, M.P. & Van Niel, K.P. (2011) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography,* **38,** 1-8.

Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberon, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling,* **222,** 1810-1819.

Baselga, A., Lobo, J.M., Svenning, J.-C. & Araújo, M.B. (2012) Global patterns in the shape of species geographical ranges reveal range determinants. *Journal of Biogeography,* **39,** 760-771.

Boakes, E.H., McGowan, P.J.K., Fuller, R.A., Chang-qing, D., Clark, N.E., O'Connor, K. & Mace, G.M. (2010) Distorted Views of Biodiversity: Spatial and Temporal Bias in Species Occurrence Data. *PLoS Biol,* **8,** e1000385.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

Dormann, C.F., Gruber, B., Winter, M. & Herrmann, D. (2010) Evolution of climate niches in European mammals? *Biology Letters,* **6,** 229-232.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution,* **1,** 330-342.

Fitzpatrick, M.C. & Hargrove, W.W. (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation,* **18,** 2255-2261.

Gallien, L., Douzet, R., Pratte, S., Zimmermann, N.E. & Thuiller, W. (2012) Invasive species distribution models - how violating the equilibrium assumption can create new insights? *Global Ecology and Biogeography*.

Heikkinen, R.K., Luoto, M., Virkkala, R., Pearson, R.G. & Korber, J.H. (2007) Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography,* **16,** 754-763.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology,* **25,** 1965-1978.

Hirzel, A., Hausser, J. & Perrin, N. (2007) Biomapper 4.0. *Laboratory for Conservation Biology, Department of Ecology and Evolution, University of Lausanne, Switzerland. URL: http://www2. unil. ch/biomapper*.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling,* **199,** 142-152.

Hulme, P.E. (2009) Trade, transport and trouble: managing invasive species pathways in an era of globalization. *Journal of Applied Ecology,* **46,** 10-18.

Hulme, P.E. & Barrett, S.C.H. (2013) Integrating trait- and niche-based approaches to assess contemporary evolution in alien plant species. *Journal of Ecology,* **101,** 68-77.

Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters,* **12,** 334-350.

Keil, P., Belmaker, J., Wilson, A.M., Unitt, P. & Jetz, W. (2013) Downscaling of species distribution models: a hierarchical approach. *Methods in Ecology and Evolution,* **4,** 82-94.

Kissling, W.D., Dormann, C.F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G.J., Montoya, J.M., Römermann, C., Schiffers, K., Schurr, F.M., Singer, A., Svenning, J.-C., Zimmermann, N.E. & O'Hara, R.B. (2012) Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography,* **39,** 2163-2178.

Koop, A., Fowler, L., Newton, L. & Caton, B. (2012) Development and validation of a weed screening tool for the United States. Biological Invasions, 14, 273-294.

McPherson, J.M., Jetz, W. & Rogers, D.J. (2006) Using coarse-grained occurrence data to predict species distributions at finer spatial resolutionsâ€"possibilities and limitations. *Ecological Modelling,* **192,** 499-522.

Menke, S.B., Holway, D.A., Fisher, R.N. & Jetz, W. (2009) Characterizing and predicting species distributions across environments and scales: Argentine ant occurrences in the eye of the beholder. *Global Ecology and Biogeography,* **18,** 50-63.

Moles, A.T., Gruber, M.A.M. & Bonser, S.P. (2008) A new framework for predicting invasive plant species. *Journal of Ecology,* **96,** 13-17.

Phillips, S.J. & Elith, J. (2010) POC plots: calibrating species distribution models with presence-only data. *Ecology,* **91,** 2476-2484.

Prinzing, A., Durka, W., Klotz, S. & Brandl, R. (2002) Geographic variability of ecological niches of plant species: are competition and stress relevant? *Ecography,* **25,** 721-729.

Rödder, D. & Engler, J.O. (2011) Quantitative metrics of overlaps in Grinnellian niches: advances and possible drawbacks. *Global Ecology and Biogeography,* **20,** 915-927.

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S.D., Grosberg, R.K., HastingS, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I. & Rice, W.R. (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution,* **22,** 465-471.

Silvertown, J., Dodd, M., Gowing, D., Lawson, C. & McConway, K. (2006) Phylogeny and the hierarchical organization of plant diversity. *Ecology,* **87,** S39-S49.

Soberón, J. & Nakamura, M. (2009) Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America,* **106,** 19644-19650.

Strubbe, D., Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography,* **22,** 962-970.

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

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental Niche Equivalency Versus Conservatism: Quantitative Approaches to Niche Evolution. *Evolution,* **62,** 2868-2883.

Whittaker, R.H. (1975) *Communities and ecosystems.* MacMillan, New York, New York, USA.

Zurell, D., Berger, U., Cabral, J.S., Jeltsch, F., Meynard, C.N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B. & Grimm, V. (2010) The virtual ecologist approach: simulating data and observers. *Oikos,* **119,** 622-635.

APPENDICES

PRESENTATION OF THE APPENDICES

During my PhD, I had the chance to be associated to interesting projects intimately related to this thesis. Here, I present briefly their aims and my contribution to each of them.

- Appendix I

This paper was the pivotal tool I used and developed to further investigate the component of niche changes in biological invaders. Among the numerous methods tested in this paper, I developed the code to include the Ecological Niche Factors Analysis (ENFA). I also assisted in the writing of the manuscript.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography,* **21,** 481-497.

- Appendix II

While I used to work with numerous species, this study focus on one remarkable species (*Centaurea stoebe*), pushing the limit of ENM to understand ecological and evolutionary processes. The colonization by this invasive species was reconstructed through both geographical and environmental space of the United States to understand when and where the niche shift processes occurred. I participated in the thinking and the writing process.

Broennimann O. , Mráz P. , Petitpierre B. , Guisan A. , Müller-Schärer H. (2014) . Contrasting spatio-temporal climatic niche dynamics during the eastern and western invasions of spotted knapweed in North America. *Journal of Biogeography*. peer-reviewed

- Appendix III

I was contacted to apply the same technique I developed in Chapter (1.2) to invasive Bryophytes. Additionally, the dimension of the extent used to calibrate ENM was taken into account. I supervised the analyses and participated to the writing. The draft will be submitted in *Journal of Biogeography*

Mateo, R. G., Broennimann, O., Petitpierre, B., Muñoz, J., van Rooy, J., Laenen, B., Guisan, A., Vanderpoorten, A. From climatic niche conservatism to spatial predictions: what can invasive bryophytes tell us?

- Appendix IV

I collaborated with statistician scientists at the EPFL to develop an *in silico* framework allowing testing different biases inherent to ENM. I provided some data and participated to the writing. The draft is currently in revision for *Methods in Ecology and Evolution*.

Thibaud, E., Petitpierre, B., Broennimann, O., Davison, A., Guisan, A. Measuring the relative effect of factors affecting species distribution model predictions

- Appendix V

I was contacted to apply the same framework used for Chapter 1.2 with introduced Amphibians and Reptiles. Additionally, we investigated for possible factors explaining niche shifts. I supervised the ENM analyses and took part to the writing process.

Li, Y., Liu, X., Li, X., Petitpierre, B. Guisan, A. (2014) Residence time, expansion towards equator in invaded range and native range size matter to climatic niche shifts among non-native species. *Global Ecology and Biogeography.* peer-reviewed.

- Appendix VI

This study is different from the main focus of my thesis but illustrates how the niche overlap problematic goes beyond the Anthropocene. I used the niche overlap framework to reconstruct the niche during the recolonization of European tree species. I was in charge to run the overlap analyses and was involved in the writing process.

Maiorano, L., Cheddadi, R., Zimmermann, N.E., Pellissier, L., Petitpierre, B., Pottier, J., Laborde, H., Hurdu, B.I., Pearman, P.B., Psomas, A., Singarayer, J.S., Broennimann, O., Vittoz, P., Dubuis, A., Edwards, M.E., Binney, H.A. & Guisan, A. (2012) Building the niche through time: using 13,000 years of data to predict the effects of climate change on three tree species in Europe. *Global Ecology and Biogeography,* **22,** 302-317.

Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2012) **21**, 481–497

A Journal of Macroecology

Global Ecology and Biogeography

RESEARCH PAPER

Measuring ecological niche overlap from occurrence and spatial environmental data

Olivier Broennimann[1]*,†, Matthew C. Fitzpatrick[2]†, Peter B. Pearman[3]†, Blaise Petitpierre[1], Loïc Pellissier[1], Nigel G. Yoccoz[4], Wilfried Thuiller[5], Marie-Josée Fortin[6], Christophe Randin[7], Niklaus E. Zimmermann[3], Catherine H. Graham[8] and Antoine Guisan[1]

[1]*Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland,* [2]*University of Maryland Center for Environmental Science, Appalachian Lab, Frostburg, MD 21532, USA,* [3]*Swiss Federal Research Institute WSL, 8903 Birmensdorf, Switzerland,* [4]*Department of Arctic and Marine Biology, University of Tromsø, 9037 Tromsø, Norway,* [5]*Laboratoire d'Ecologie Alpine, University Joseph Fourier, 38041 Grenoble, France,* [6]*Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Canada M5S 1A1,* [7]*Institute of Botany, University of Basel, 4056 Basel, Switzerland,* [8]*Department of Ecology and Evolution, SUNY at Stony Brook, NY 11794, USA*

*Correspondence: Olivier Broennimann, Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland.
E-mail: olivier.broennimann@unil.ch
†The first three authors have contributed equally to this paper.

ABSTRACT

Aim Concerns over how global change will influence species distributions, in conjunction with increased emphasis on understanding niche dynamics in evolutionary and community contexts, highlight the growing need for robust methods to quantify niche differences between or within taxa. We propose a statistical framework to describe and compare environmental niches from occurrence and spatial environmental data.

Location Europe, North America and South America.

Methods The framework applies kernel smoothers to densities of species occurrence in gridded environmental space to calculate metrics of niche overlap and test hypotheses regarding niche conservatism. We use this framework and simulated species with pre-defined distributions and amounts of niche overlap to evaluate several ordination and species distribution modelling techniques for quantifying niche overlap. We illustrate the approach with data on two well-studied invasive species.

Results We show that niche overlap can be accurately detected with the framework when variables driving the distributions are known. The method is robust to known and previously undocumented biases related to the dependence of species occurrences on the frequency of environmental conditions that occur across geographical space. The use of a kernel smoother makes the process of moving from geographical space to multivariate environmental space independent of both sampling effort and arbitrary choice of resolution in environmental space. However, the use of ordination and species distribution model techniques for selecting, combining and weighting variables on which niche overlap is calculated provide contrasting results.

Main conclusions The framework meets the increasing need for robust methods to quantify niche differences. It is appropriate for studying niche differences between species, subspecies or intra-specific lineages that differ in their geographical distributions. Alternatively, it can be used to measure the degree to which the environmental niche of a species or intra-specific lineage has changed over time.

Keywords

***Centaurea stoebe*, ecological niche model, kernel density, niche conservatism, niche equivalency, niche similarity, ordination, *Solenopsis invicta*, species distribution model, virtual species.**

© 2011 Blackwell Publishing Ltd

DOI: 10.1111/j.1466-8238.2011.00698.x
http://wileyonlinelibrary.com/journal/geb

INTRODUCTION

> It is, of course, axiomatic that no two species regularly established in a single fauna have precisely the same niche relationships.
>
> (Grinnell, 1917)

An ongoing challenge for ecologists is quantifying species distributions and determining which factors influence species range limits (Guisan & Thuiller, 2005; Colwell & Rangel, 2009). Factors that can constrain species distributions include abiotic gradients, such as climate, sunlight, topography and soils, and biotic interactions, such as the identity and abundance of facilitators (e.g. pollinators, seed dispersers), predators, parasites and competitors (Gaston, 2003). The study of how species vary in their requirements for and tolerance of these factors has advanced, in part due to the continued conceptual development and quantification of the ecological niche of species (Chase & Leibold, 2003; Soberón, 2007). The complementary concepts of the environmental niche (*sensu* Grinnell, 1917) and the trophic niche (*sensu* Elton, 1927) serve as a basis for assessing the ecological and biogeographical similarities and differences among species. Toward this end, a variety of measures have been used to quantify niche characteristics. Historically, such assessments have focused primarily on differences in local trophic and reproductive habits (reviewed in Chase & Leibold, 2003) and have asked: How much does resource use by species A overlap with that of species B? Recent concern over the effects of global change on species distributions has emphasized the need to quantify differences among species in their environmental requirements in a geographical context and at an extent comparable to that of species ranges. Consistent with aspects of the Grinnellian niche, such assessments pursue questions regarding similarities and differences in the environmental conditions associated with species geographical distributions and how they change over time (Devictor *et al*., 2010). Despite improvements in our ability to model species distributions (Guisan & Thuiller, 2005), the development of techniques to quantify overlap of different environmental niches has received relatively little attention (but see Warren *et al*., 2008).

A variety of approaches and metrics have been used to measure niche overlap (e.g. Horn, 1966; MacArthur & Levins, 1967; Schoener, 1970; Colwell & Futuyma, 1971; May & Arthur, 1972; Pianka, 1980). Generally, these methods date to the period in which competition was widely believed to be the primary mechanism structuring ecological communities and measures of niche overlap were developed to quantify differences due to competition (Chase & Leibold, 2003). More recently, research has elucidated how changing environmental conditions could affect future distributions of native species (e.g. Etterson & Shaw, 2001; Jump & Peñuelas, 2005) and invasive exotic species (e.g. Broennimann *et al*., 2007; Fitzpatrick *et al*., 2007; Steiner *et al*., 2008; Medley, 2010). Further, changes in the climatic tolerances and requirements of species accompany the diversification of lineages in a variety of taxa (e.g. Silvertown *et al*., 2001; Losos *et al*., 2003; Yesson & Culham, 2006; Evans *et al*., 2009). A common theme among these studies is the quantification of environmental niches, how they change over time and differ among species. Yet the inadequacy of methods for comparing species environmental niches has fuelled debate over the validity of conclusions derived from comparative studies of niche dynamics (Fitzpatrick *et al*., 2008; Peterson & Nakazawa, 2008).

Assessing differences in the environmental niches of species requires identification and consideration of the factors that influence species distributions. In practice, distributions of species are often characterized using occurrence records (Graham *et al*., 2004). Differences in niches that are quantified using observed occurrences of species reflect an unknown conjunction of the environmental niches of the species, the biotic interactions they experience and the habitats available to species and colonized by them (Soberón, 2007; Colwell & Rangel, 2009). Although it has often been assumed that these effects are negligible at broad spatial scales, recent studies indicate that biotic interactions may play an important role in defining the lower thermal boundaries of species distributions (e.g. Gotelli *et al*., 2010; Sunday *et al*., 2011). This subset of the environmental niche that is actually occupied by the species corresponds to the realized niche (Hutchinson, 1957). The environmental conditions comprising the realized niche are described using a set of geographically referenced environmental variables. These variables come from widely used, on-line collections such as WorldClim (Hijmans *et al*., 2005), a wealth of other variables of some physiological and demographic importance (e.g. Zimmermann *et al*., 2009), and physical habitat variation as represented by country and regional land cover as well as land-use classifications (e.g. Lütolf *et al*., 2009). Hereafter, the use of geographically referenced variables is often implicit when we refer to niche comparison, but the approaches and metrics we present can be applied to any quantitative niche dimension.

Methods for quantifying the environmental niche and estimating niche differences typically rely on either ordination techniques (e.g. Thuiller *et al*., 2005a; Hof *et al*., 2010) or species distribution models (SDMs; Guisan & Thuiller, 2005) Ordination techniques allow for direct comparisons of species–environment relationships in environmental space, and employ various maximization criteria to construct synthetic axes from associated environmental variables (Jongman *et al*., 1995). In contrast, assessment of niche differences with SDMs involves calibration (for each species) of statistical or machine-learning functions that relate environmental variables to georeferenced data on species occurrence (Guisan & Thuiller, 2005). SDMs can select and emphasize, via weighting, certain variables associated with processes that determine the distribution of the species (through their environmental niches) while down-weighting or excluding variables that do not help to discriminate between species presence and absence (Wintle *et al*., 2003; Guisan & Thuiller, 2005). Niche overlap is then estimated through the projection of those functions across a landscape (i.e. the overlap is calculated in geographical space). Recently, Warren *et al*. (2008) developed such a SDM-based method that uses cell-by-cell comparisons of geographical predictions of occurrences and randomization tests to quantify niche differences and assess

their statistical significance. However, niche overlap analyses using geographical projections of niches derived from SDMs could prove problematic because the measured niche overlap is likely to vary depending on the extent and distribution of environmental gradients in the study area and potentially because of unquantified statistical artefacts related to model fitting.

Here, we present a new statistical framework to describe and compare niches in a gridded environmental space (i.e. where each cell corresponds to a unique set of environmental conditions). Within this framework, we quantify niche overlap using several ordination and SDM techniques and evaluate their performance. The framework overcomes some of the shortcomings of current approaches to quantifying niche differences: (1) it accounts for biases introduced by spatial resolution (grid size), (2) makes optimal use of both geographical and environmental spaces, and (3) corrects observed occurrence densities for each region in light of the availability of environmental space. Case studies from nature are unlikely to provide an unbiased assessment of methods used to quantify niche overlap because of sampling errors and unknown biases. To overcome these issues, we test the methods using simulated species distributions for which niche overlap and the constraining environmental gradients are known without error. Finally, we illustrate our approach using two invasive species that have native and invaded ranges on different continents and which have been subjects of recent studies of niche dynamics (Broennimann *et al.*, 2007; Fitzpatrick *et al.*, 2007).

METHODS

A framework to compare environmental niches

We present a framework for quantifying niche overlap between two species (e.g. sister taxa, subspecies, etc.) or two distinct sets of populations of a same species (e.g. native and invasive populations of an invasive species, geographically disjunct populations of the same species, etc.). The framework also applies to comparisons among the same species but at different times (e.g. before and after climate change). More broadly, the framework can be applied to compare any taxonomic, geographical or temporal groups of occurrences (hereafter called 'entities'). The framework involves three steps: (1) calculation of the density of occurrences and of environmental factors along the environmental axes of a multivariate analysis (2) measurement of niche overlap along the gradients of this multivariate analysis and (3) statistical tests of niche equivalency and similarity (*cf.* Warren *et al.*, 2008). All the analyses are done in R (R Development Core Team, 2010) and scripts are available online in Appendix S1 in the Supporting Information.

Calibration of the niche and occurrence density

The environmental space is defined by the axes of the chosen analysis and is bounded by the minimum and maximum environmental values found across the entire study region. In this application, we consider the first two axes for ordinations such as principal components analysis (PCA) and one axis for SDMs (i.e. the output of a SDM comprises a single vector of predicted probabilities of occurrence derived from complex combinations of functions of original environmental variables; the overlap of the two species is analysed along this gradient of predictions). We recognize that, in principle, niche overlap analyses can consider greater dimensionality than we do here. However, in practice increased dimensionality brings greater challenges in terms of interpretation, visualization and additional technical aspects. Nonetheless, a greater number of dimensions should be considered in further development of the present approach. The environmental space is divided into a grid of $r \times r$ cells (or a vector of r-values when the analysis considers only one axis). For our analyses we set the resolution r to 100. Each cell corresponds to a unique vector of environmental conditions v_{ij} present at one or more sites in geographical space, where 'i' and 'j' refer to the cell corresponding respectively to the ith and jth bins of the environmental variables. The bins are defined by the chosen resolution r, and the minimum and maximum values present in the study area along these variables.

Since the number of occurrences is dependent on sampling strategy, sampled occurrences may not represent the entire distribution of the species or other taxon nor the entire range of suitable environmental conditions, resulting in underestimated densities in some cells and potentially large bias in measured niche overlap (Fig. S1a). Interestingly, this problem is similar to the delimitation of the utilization distribution of species in geographical space. Traditionally, methods such as minimum convex polygons have been used to delimitate utilization distributions (e.g. Blair, 1940). But, newer developments have shown that kernel methods provide more informative estimations (Worton, 1989), and such methods have seen recent application in modelling species distributions (Ferrier *et al.*, 2007; Hengl *et al.*, 2009). We thus apply a kernel density function to determine the 'smoothed' density of occurrences in each cell in the environmental space for each dataset. We use the standard smoothing parameters used in most density estimation studies (Gaussian kernel with a standard bandwidth, which corresponds to 0.9 times the minimum of the standard deviation and the inter-quartile range of the data divided by 1.34 times the sample size to the negative one-fifth power; Silverman, 1986). The smoothed density of occurrence o_{ij} for each cell is thus calculated as

$$o_{ij} = \frac{\delta(n_{ij})}{\max(n_{ij})}, \quad (1)$$

where $\delta(n_{ij})$ is the kernel density estimation of the number of occurrences of the entity at sites with environment v_{ij}, $\max(n_{ij})$ is the maximum number of occurrences in any one cell, and o_{ij} is a relative abundance index that ranges from 0 for environmental conditions in which the entity has not been observed, to 1 for environmental conditions in which the entity was most commonly observed. In a similar manner, the smoothed density of available environments e_{ij} is calculated as

$$e_{ij} = \frac{\delta(N_{ij})}{\max(N_{ij})}, \quad (2)$$

where $\delta(N_{ij})$ is the number of sites with environment v_{ij} and $\max(N_{ij})$ is the number of cells with the most common environment in the study area. Finally, we calculate z_{ij}, the occupancy of the environment v_{ij} by the entity, as

$$z_{ij} = \frac{o_{ij}/e_{ij}}{\max(o/e)} \quad \text{if} \quad e_{ij} \neq 0 \quad (3a)$$

and

$$z_{ij} = 0 \quad \text{if} \quad e_{ij} = 0, \quad (3b)$$

where z_{ij} ranges between 0 and 1 and ensures a direct and unbiased comparison of occurrence densities between different entities occurring in ranges where environments are not equally available.

Measurement of niche overlap

The comparison of z_{ij} between two entities can be used to calculate niche overlap using the *D* metric (Schoener, 1970; reviewed in Warren *et al*., 2008) as

$$D = 1 - \frac{1}{2}\left(\sum_{ij} |z_{1ij} - z_{2ij}|\right),$$

where z_{1ij} is entity 1 occupancy and z_{2ij} is entity 2 occupancy. This metric varies between 0 (no overlap) and 1 (complete overlap). Note that regions of the environmental space that do not exist in geography have z_{ij} set to 0. These regions thus do not contribute to the measure of the *D* metric and niche overlap is measured among real habitats only (see discussion in Warren *et al*., 2008). Note also that the use of a kernel density function when calculating the density is critical for an unbiased estimate of *D*. When no kernel density function is applied, the calculated overlap depends on the resolution *r* chosen for the gridded environmental space (Fig. S1a). Using smoothed densities from a kernel density function ensures that the measured overlap is independent of the resolution of the grid (Fig. S1b).

Statistical tests of niche equivalency and similarity

We built from the methodology described in Warren *et al.* (2008) to perform niche equivalency and similarity tests. The niche equivalency test determines whether niches of two entities in two geographical ranges are equivalent (i.e. whether the niche overlap is constant when randomly reallocating the occurrences of both entities among the two ranges). All occurrences are pooled and randomly split into two datasets, maintaining the number of occurrences as in the original datasets, and the niche overlap statistic *D* is calculated. This process is repeated 100 times (to ensure that the null hypothesis can be rejected with high confidence) and a histogram of simulated values is constructed. If the observed value of *D* falls within the density of 95% of the simulated values, the null hypothesis of niche equivalency cannot be rejected.

The niche similarity test differs from the equivalency test because the former examines whether the overlap between observed niches in two ranges is different from the overlap between the observed niche in one range and niches selected at random from the other range. In other words, the niche similarity test addresses whether the environmental niche occupied in one range is more similar to the one occupied in the other range than would be expected by chance. For this test, we randomly shift the entire observed density of occurrences in one range (the centre of the simulated density of occurrence is randomly picked among available environments) and calculate the overlap of the simulated niche with the observed niche in the other range. The test of niche similarity is also based on 100 repetitions. If the observed overlap is greater than 95% of the simulated values, the entity occupies environments in both of its ranges that are more similar to each other than expected by chance. Note that in some instances it may be difficult to define the extent of the study areas to be compared. When species occur on different continents, the choice can be straightforward and should consider the complete gradient of environmental space that the study species could reasonably encounter, including consideration of dispersal ability and major biogeographical barriers or transitions. When species occur in the same region or on an island, the environment can be the same for all species and therefore correcting for differences in the densities of environments is not necessary.

Testing the framework with virtual entities

A robust test of the framework described above requires entities that have distributions determined by known environmental parameters and that exhibit known levels of niche overlap. To achieve this, we simulated pairs of virtual entities with varying amounts of niche overlap (Fig. 1, see also Appendix S2 for details), in a study region comprising all temperate climates in Europe (EU) and North America (NA) and defined by eight bioclimatic variables at 10′ resolution that were derived from raw climatic data from the CRU CL 2.0 dataset (New *et al*., 2002). These variables included: ratio of actual and potential evapotranspiration (aetpet), number of growing degree days above 5 °C (gdd), annual precipitation (p), potential evapotranspiration (pet), number of months with drought (ppi), seasonality in precipitation (stdp), annual mean temperature (t), annual maximum temperature (tmax), and annual minimum temperature (tmin). Procedures to calculate aetpet, pet and gdd from the raw CRU CL 2.0 data are detailed in Thuiller *et al.* (2005b).

We first apply the framework to 100 pairs of virtual entities that differ in niche position and that exhibit decreasing amounts of niche overlap, from perfect overlap ($D = 1$, all areas in common under the normal density curves) to no overlap ($D = 0$, no area in common under the normal density curves). We compare these simulated levels of niche overlap to that mea-

Table 1 Ordination techniques for quantifying niche overlap. In addition to a general description of the technique, an explanation of its application to the comparison of simulated niches between the European (EU) and North American (NA) continents is provided. Depending on the type of analysis and whether a priori groups are used or not, the different areas of calibration we tested are specified.

Name	Description	Areas of calibration
PCA-occ	Principal component analysis (Pearson, 1901) transforms a number of correlated variables into a small number of uncorrelated linear combinations of the original variables (principal components). These components are the best predictors – in terms of R^2 – of the original variables. In other terms, the first principal component accounts for as much of the variability in the data as possible, and each following component accounts for as much of the remaining variability as possible. For the study of niche overlap, the data used to calibrate the PCA are the climate values associated with the occurrences of the species. Additional occurrence data can be projected in the same environmental space. When calibrating the PCA with EU and NA occurrences, differences in position along the principal components discriminate environmental differences between the two distributions. When calibrating with EU occurrences only, differences in position along the principal components maximize the discrimination of differences among the EU distribution	1. Occ. in EU 2. Occ. in EU+NA
PCA-env	Same as PCA-occ but calibrated on the entire environmental space of the two study areas, including species occurrences. When calibrating PCA-env on EU and NA ranges, differences in position along the principal components discriminate differences between the EU and NA environmental spaces whereas a calibration on the EU full environmental space maximizes the discrimination among this range only	1. EU range 2. EU&NA ranges
BETWEEN-occ and WITHIN-occ	Between-group and within-group analyses (Dolédec & Chessel, 1987) are two ordination techniques that rely on a primary analysis (here PCA, but could be CA or MCA) but use a priori groups to optimize the combination of variable in the principal components. Here the a priori groups correspond to EU and NA. BETWEEN-occ and WITHIN-occ are calibrated with EU&NA occurrences, and respectively maximize or minimize the discrimination of niche differences between EU and NA occurrences	1. Occ. in EU+NA
WITHIN-env	Same as WITHIN-occ but calibrated on the entire environmental spaces of the two continents. WITHIN-env minimizes the discrimination of environmental differences between EU and NA ranges	1. EU&NA ranges
LDA	Linear discriminant analysis (LDA; Fisher, 1936) finds linear combinations of variables which discriminate the differences between two or more groups. The objective is thus similar to BETWEEN but uses a different algorithm. Distances between occurrences are calculated with Mahalanobis distance	1. Occ. in EU+NA
MDS	Multidimensional scaling (MDS; Gower, 1966) is a nonparametric generalization of PCA that allows various choices of measures of associations (not limited to correlation and covariance as in PCA). Here we use the distance in the Euclidean space. The degree of correspondence between the distances among points implied by MDS plot and the input distance structure is measured (inversely) by a *stress* function. Scores are juggled to reduce the stress until stress is stabilized	1. Occ. in EU 2. Occ. in EU+NA
ENFA	Ecological niche factor analysis (ENFA; Hirzel *et al.*, 2002). ENFA is an ordination technique that compares environmental variation in the species distribution to the entire area. This method differs from other ordination techniques in that the principal components have a direct ecological interpretation. The first component corresponds to a marginality factor: the axis on which the species niche differs at most from the available conditions in the entire area. The next components correspond to specialization factors: axes that maximize the ratio of the variance of the global distribution to that of the species distribution	1. Occ. in EU + EU range 2. Occ. in EU&NA + EU&NA ranges

CA, correspondence analysis; MCA, multiple correspondence analysis.

sured along the p and t gradients (instead of the two first axes of a multivariate analysis). Since the normal density curves defining the niches of the virtual entities (Appendix S2) are built along these two gradients, we postulate that the overlap detected by the application of the framework should be the same as the simulated level of niche overlap across the full range of possible overlaps (0 to 1).

Next, we apply the framework to pairs of virtual entities but compare the simulated level of niche overlap with the niche overlap detected along axes calibrated using several ordination (Table 1) and SDM techniques (Table 2). For methods with maximization criteria that do not depend on an a priori grouping (here EU versus NA, Table 1), we run two sets of simulations, using information from either EU alone or both EU and NA to

Table 2 Species distribution modelling (SDM) techniques for quantifying niche overlap. GLM, GBM and RF were fitted with species presence–absence as the response variable and environmental variables as predictors (i.e. explanatory variables) using the BIOMOD package in R (Thuiller *et al*., 2009, R-Forge.R-project.org) and default settings. MaxEnt was fitted using the *dismo* package in R with default settings. For all techniques, we use pseudo-absences that were generated randomly throughout the area of calibration. Two sets of models were created using two areas of calibration: one using presence–absence data in Europe (EU) only and a second using presence-absence data in both EU and North America (NA). The resulting predictions of occurrence of the species (ranging between 0 and 1) are used as environmental axes in the niche overlap framework.

Name	Description
GLM	Generalized linear models (GLM; McCullagh & Nelder, 1989) constitute a flexible family of regression models, which allow several distributions for the response variable and non-constant variance functions to be modelled. Here we use binomial (presence–absence) response variables with a logit link function (logistic regression) and allow linear and quadratic relationship between the response and explanatory variables. A stepwise procedure in both directions was used for predictor selection, based on the Akaike information criterion (AIC; Akaike, 1974).
MaxEnt	MaxEnt (Phillips *et al*., 2006) is a machine learning algorithm that estimates the probability of occurrence of a species in contrast to the background environmental conditions. MaxEnt estimates species distributions by finding the distribution of maximum entropy (i.e. that is most spread out, or closest to uniform) subject to the constraint that the expected value for each environmental variable under this estimated distribution matches its empirical average. MaxEnt begins with a uniform distribution then uses an iterative approach to increase the probability value over locations with conditions similar to samples. The probability increases iteration by iteration, until the change from one iteration to the next falls below the convergence threshold. MaxEnt uses $L - 1$ regularization as an alternative to stepwise model selection to find parsimonious models
GBM	The gradient boosting machine (GBM; Friedman, 2001) is an iterative computer learning algorithm. In GBM, model fitting occurs not in parameter space but instead in function space. The GBM iteratively fits shallow regression trees, updating a base function with additional regression tree models. A randomly chosen part of the training data is used for function fitting, leaving the other part for estimating the optimal number of trees to use during prediction with the model (out-of-bag estimate)
RF	Random forests (RF; Breiman, 2001). Random forests grows many classification trees. To classify the species observations (i.e. presences and absences) from the environmental variables, RFs puts the variables down each of the trees in the forest. Each tree gives a classification, and the tree 'votes' for that class. The forest chooses the classification having the most votes (over all the trees in the forest). Random forests is designed to avoid overfitting

calibrate the method ('Areas of Calibration', Tables 1 & 2). To compare the outcomes of the methods quantitatively, for each analysis we first calculate the average absolute difference between the simulated and measured overlap (Δ_{abs}). This difference indicates the magnitude of the errors (deviation from the simulated = measured diagonal). To test for biases in the method (i.e. whether or not scores are centred on the diagonal), we then perform a Wilcoxon signed-rank test on these differences. A method that reliably measures simulated levels of niche overlap should show both small errors (small Δ_{abs}) and low bias (non-significant Wilcoxon test).

Case studies for real species

We also test the framework using two invasive species that have native and invaded ranges on different continents and which have been subjects of recent analyses of niche dynamics. The first case study concerns spotted knapweed (*Centaurea stoebe*, Asteraceae), native to Europe and highly invasive in North America (see Broennimann *et al*., 2007; Broennimann & Guisan, 2008 for details). The second case study addresses the fire ant (*Solenopsis invicta*), native to South America and invasive in the USA (see Fitzpatrick *et al*., 2007, 2008 for details).

RESULTS

Evaluation of the framework

Before applying ordination and SDM methods to our datasets, we examined whether we could accurately measure simulated levels of niche overlap along known gradients. We used 100 pairs of virtual entities with known levels of niche overlap along p and t climate gradients. The overlap we detected between each pair of virtual entities is almost identical to the simulated overlap (i.e. the shared volume between the two simulated bivariate normal curves; filled circles, Fig. 1). This is the case for all levels of overlap except for highly overlapping distributions (> 0.8) where the actual overlap is slightly underestimated, and where the effects of sampling are likely to be most evident. Because detected overlap cannot be larger than 1, any error in the measurement of highly overlapping distribution must necessarily result in underestimation. This underestimation is, however, very small (Δ_{abs}:$\mu = 0.024$) and does not alter interpretation. Note that when overlap is measured using virtual entities that follow a univariate normal distribution along a precipitation gradient, no underestimation was observed (Fig. S2). When we leave differences in environmental availability uncorrected, niche overlap is consistently underestimated (open circles,

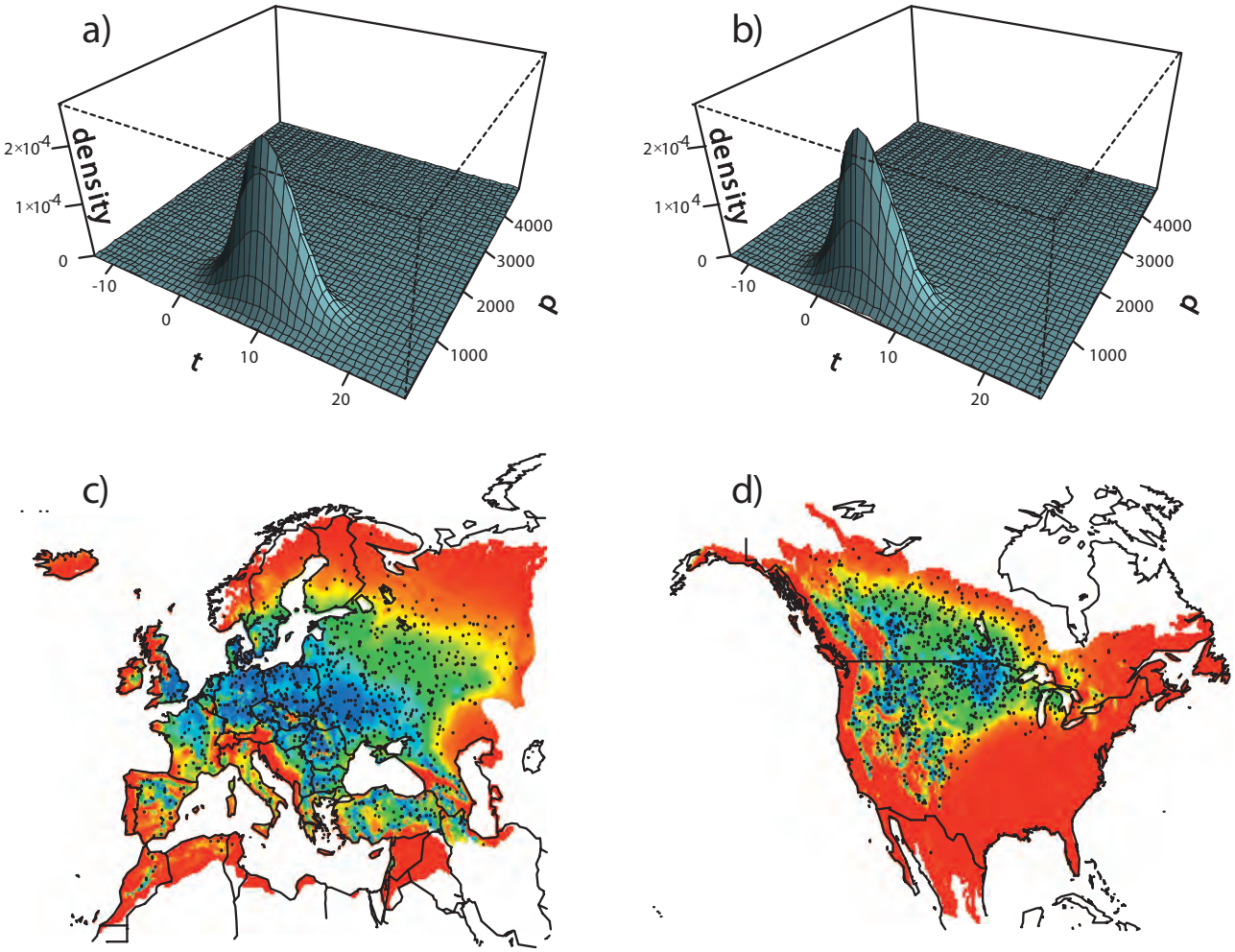


Figure 1 Example of virtual species following a bivariate normal density along precipitation (p) and temperature (t) gradients with 50% overlap between the European and North American niche in environmental space. The red to blue colour scale shows the projection of the normal densities in the geographical space from low to high probabilities (i.e. 0 to 1). Black dots show random occurrences.

Fig. 2), except for niches with low overlap (< 0.3). This bias is on average five times larger than that of the corrected measure.

Niche overlap detected by ordination and SDM methods

Simulated entities

Ordination and SDM techniques vary in their ability to measure simulated niche overlap (Figs 3–5). Among methods with maximization criteria that do not depend on a priori grouping (Fig. 3), PCA-env calibrated on both EU and NA ranges most accurately measures simulated niche overlap (Δ_{abs}:μ = 0.054, Wilcoxon $P > 0.05$; Fig. 3b). Note, however, that highly overlapping distributions are somewhat underestimated but the significance of the Wilcoxon test is unaffected. The only other predominantly unbiased method in this category is ecological niche factor analysis (ENFA), also calibrated on environmental data from both ranges. However, errors generated by ENFA are comparatively high (Δ_{abs}:μ = 0.156, Wilcoxon $P > 0.05$; Fig. 3d). Scores of PCA-occ and MDS are significantly biased, with the measured overlap consistently lower than the simulated one (Fig. 3a, b), especially in the ordination of data combined from both EU and NA ranges.

Among methods with maximization criteria based on a priori grouping (Fig. 4), WITHIN-env provides the lowest errors of measured overlap. However, WITHIN-env significantly underestimates the simulated overlap (Δ_{abs}:μ = 0.084, Wilcoxon $P < 0.001$; Fig. 4b), though the amount of underestimation is small. By contrast, WITHIN-occ overestimates simulated overlap (Δ_{abs}:μ = 0.195, Wilcoxon $P < 0.001$; Fig. 4a). Predictably, techniques that maximize discrimination between groups (BETWEEN-occ and LDA; Fig. 4c, d) fail to measure simulated levels of niche overlap adequately. Both methods provide similar results in which overlap is underestimated across all simulated levels.

Compared with ordinations, SDM methods show different patterns when measuring overlap (Fig. 5). When calibrated on both ranges, all SDM methods report high levels of overlap (0.6–1), regardless of simulated overlap. SDMs apparently cali-

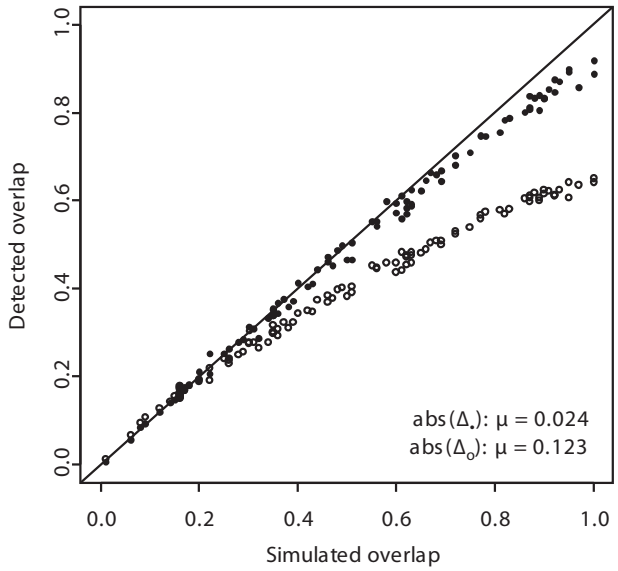


Figure 2 Agreement between simulated and detected niche overlap. Each dot corresponds to a pair of simulated entities. Simulated overlap corresponds to the volume in common between the two bivariate normal distributions with different means on precipitation and temperature gradients (see Figure 1). Filled circles represent the detected overlap with correction for climate availability (density of occurrences divided by the density of climate across the entire climate space). Open circles show the detected overlap when no correction for climate availability is applied. The average absolute difference between the simulated and measured overlap (abs(Δ):μ) is indicated for both corrected and uncorrected measures.

brate bimodal curves that tightly fit the two distributions as a whole. However, when calibrated on the EU range only, all SDM methods report increasing levels of overlap along the gradient of simulated overlap. MaxEnt achieves the best results (Δ_{abs}:μ = 0.111, Wilcoxon $P > 0.05$; Fig. 5b), followed by the gradient boosting machine (GBM) (Δ_{abs}:μ = 0.134, Wilcoxon $P < 0.05$; Fig. 5c). MaxEnt is the only SDM method providing non-significant bias. Generalized linear modelling (GLM) exhibits a similar amount of error as GBM, but with lower reported overlap (Δ_{abs}:μ = 0.147, Wilcoxon $P < 0.001$; Fig. 5a). Random forests (RF) provides very poor results in terms of both error and bias (Δ_{abs}:μ = 0.393, Wilcoxon $P < 0.001$; Fig. 5d).

Case studies

Analyses of spotted knapweed and fire ant occurrences using PCA-env, the most accurate method in terms of niche overlap detection, show that for both species the niche in the native and invaded ranges overlap little (0.25 and 0.28 respectively, Figs 6 & 7). For spotted knapweed, the invaded niche exhibits both shift and expansion (Fig. 6a, b) relative to its native range. Interestingly, two regions of dense occurrence in NA indicate two known areas of invasion in western and eastern NA. In contrast, the fire ant exhibits a shift from high density in warm and wet environments in South America towards occupying cooler and drier environments in NA (Fig. 7a, b). For both species, niche equivalency is rejected, indicating that the two species have undergone significant alteration of their environmental niche during the invasion process (Figs 6d & 7d). However, for both species, niche overlap falls within the 95% confidence limits of the null distributions, leading to non-rejection of the hypothesis of retained niche similarity (Figs 6e & 7e).

DISCUSSION

The framework we have presented helps meet the increasing need for robust methods to quantify niche differences between or within taxa (Wiens & Graham, 2005; Pearman *et al*., 2008a). By using simulated entities with known amounts of niche overlap, our results show that niche overlap can be accurately detected within this framework (Fig. 2). Our method is appropriate for the study of between-species differences of niches (e.g. Thuiller *et al*., 2005a; Hof *et al*., 2010), as well as to compare subspecies or distinct populations of the same species that differ in their geographical distributions and which are therefore likely to experience different climatic conditions (e.g. Broennimann *et al*., 2007; Fitzpatrick *et al*., 2007; Steiner *et al*., 2008; Medley, 2010). Alternatively, when a record of the distribution of the taxa (and corresponding environment) through time exists, our approach can be used to answer the question of whether and to what degree environmental niches have changed through time (e.g. Pearman *et al*., 2008b; Varela *et al*., 2010).

This framework presents two main advantages over methods developed previously. First, it disentangles the dependence of species occurrences from the frequency of different climatic conditions that occur across a region. This is accomplished by dividing the number of times a species occurs in a given environment by the frequency of locations in the region that have those environmental conditions, thereby correcting for differences in the relative availability of environments. Without this correction, the measured amount of niche overlap between two entities is systematically underestimated (Fig. 2). For example, in the approach of Warren *et al*. (2008), who used an SDM-based method using comparisons of geographical predictions of occurrences, projections depend on a given study area. Measured differences between niches could represent differences in the environmental characteristics of the study area rather than real differences between species. Second, application of a kernel smoother to standardized species densities makes moving from geographical space, where the species occur, to the multivariate environmental space, where analyses are performed, independent of both sampling effort and of the resolution in environmental space (Fig. S1). This is a critical consideration, because it is unlikely that species occurrences and environmental datasets from different geographical regions or times always present the same spatial resolution. Without accounting for these differences, measured niche overlap will partially be a function of data resolution.

Although niche overlap can be detected accurately when variables driving the distribution are known (e.g. with niches

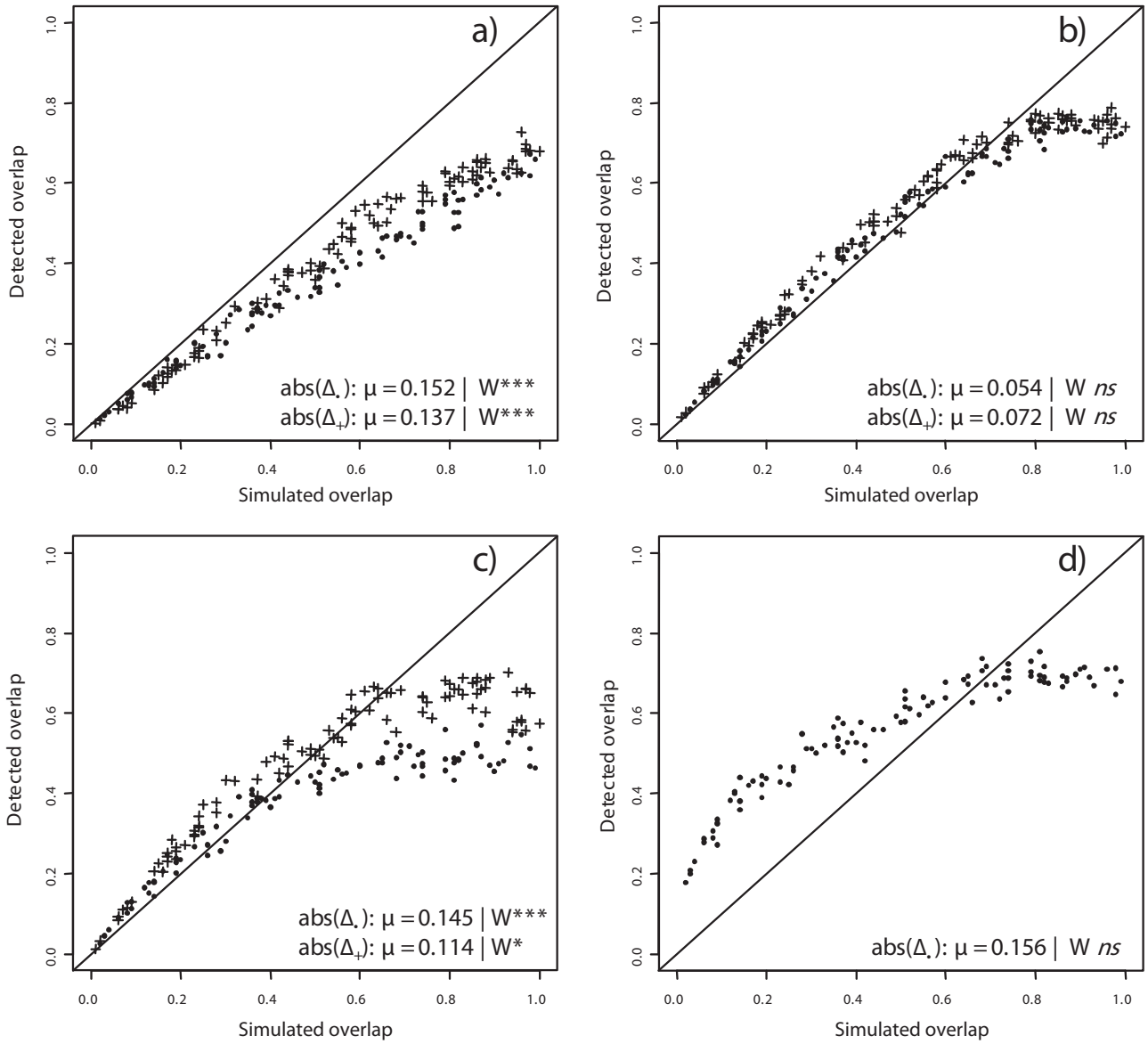


Figure 3 Sensitivity analysis of simulated versus detected niche overlap for ordinations not using a grouping factor. The axes of the analyses on which the overlap is measured correspond to (a, b) principal components analyses (a, PCA-occ; b, PCA-env), (c) multidimensional scaling (MDS) and (d) ecological niche factor analysis (ENFA). Crosses refer to models calibrated on the European (EU) range only. Black dots indicate models calibrated on both EU and North American (NA) ranges. Results for ENFA calibrated on the EU range only could not be provided because of computational limitations. Abs(Δ):μ indicate the average absolute difference between simulated and detected overlaps. The significance of the Wilcoxon signed-rank test, *W*, is shown (ns, non-significant; $^{*}0.05 < P < 0.01$; $^{***}P < 0.001$).

defined along precipitation and temperature gradients, Fig. 1), the use of ordination and SDM techniques for selecting, combining and weighting variables on which the overlap is calculated provide contrasting results. The causes of the differences in performance among techniques remain unclear, but several factors might be responsible. Among the important factors are: (1) how the environment varies in relation to species occurrences versus the study region (or time period) as a whole, (2) how techniques select variables based on this variation, and (3) the level of collinearity that exists between variables within each area/time and whether it remains constant among areas/times. Hereafter we discuss the performance of the techniques we tested in the light of these factors.

Ordinations versus SDMs

Ordinations and SDMs use contrasting approaches to reduce the dimensions of an environmental dataset. While ordinations find orthogonal and linear combinations of original predictors that maximize a particular ratio of environmental variance in

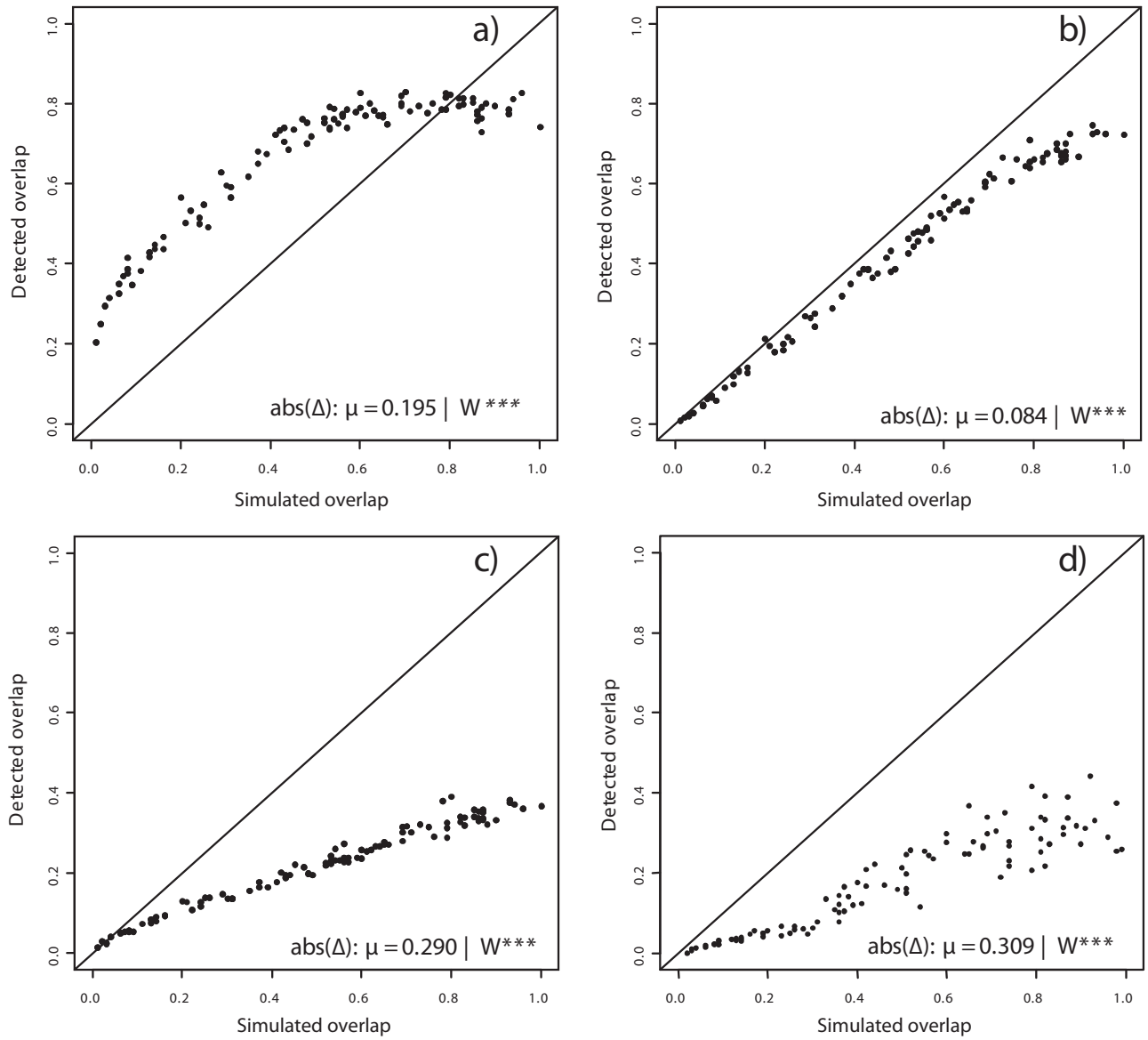


Figure 4 Sensitivity analysis of simulated versus detected niche overlap for ordinations using a priori grouping factor. The axes of the analyses on which the overlap is measured correspond to (a, b) within-group analyses (a, WITHIN-occ; b, WITHIN-env), (c) between-group analysis (BETWEEN-occ) and (d) linear discriminant analysis (LDA). Black dots indicate models calibrated on both EU and NA ranges. Abs(Δ):μ indicates the average absolute difference between simulated and detected overlaps. The significance of the Wilcoxon signed-rank test, *W*, is shown (***$P < 0.001$).

the dataset, SDMs fit nonlinear response curves, attributing different weights to variables according to their capacity to discriminate presences from absences (or pseudo-absences). When using both study regions for the calibration, SDMs consistently overestimate the simulated level of niche overlap (Fig. 5, black circles). It is likely that SDMs fit bimodal response curves that tightly match the data and artificially predict occurrences in both ranges (i.e. SDMs model the range of each entity as a single complex, albeit overfitted, niche). As a result, prediction values for occurrences are high for both ranges. Since the overlap is measured on the gradient of predicted values, measured overlap is inevitably high. In contrast, ordinations calibrated on both areas provide a simpler environmental space (i.e. a linear combination of original predictors), in which niche differences are conserved. As a result, ordinations usually show a monotonic relationship between detected and simulated overlap (Figs 3 & 4, black circles).

When calibrating SDMs using only one study area and subsequently projecting the model to another area, estimated overlap increases with simulated overlap (Fig. 5, crosses). However, the pattern of detected overlap using SDMs is irregular (i.e. Δ_{abs}:μ is high), again probably because of overfitting. Bias in detected overlap may also arise from the differing spatial structure of environments between study areas. Unlike ordina-

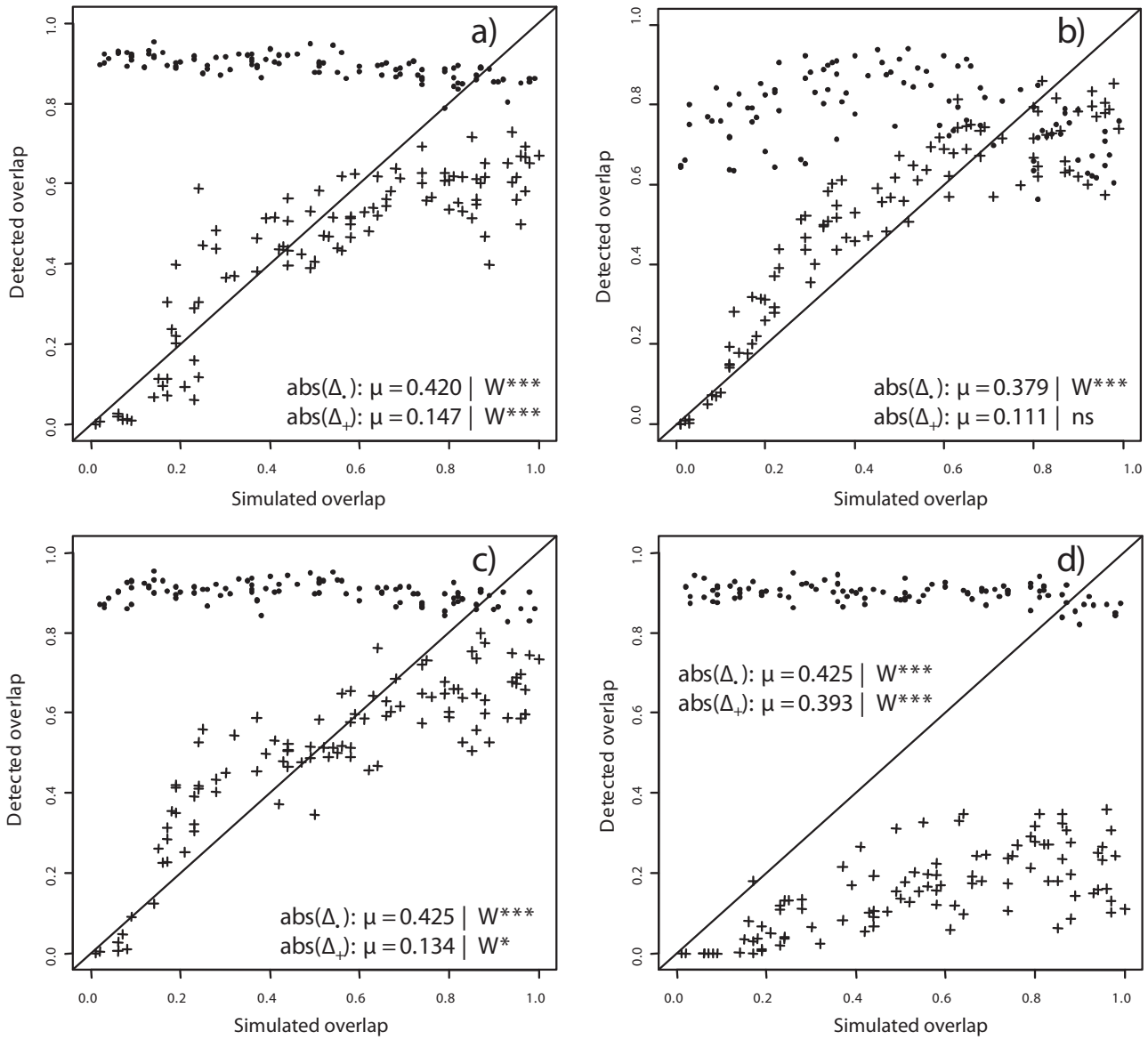


Figure 5 Sensitivity analysis of simulated versus detected niche overlap for different species distribution model (SDM) algorithms. The axes of the analyses on which the overlap is measured correspond to (a) generalized linear models (GLM), (b) MaxEnt, (c) gradient boosting machine (GBM) and (d) random forests (RF). Crosses refer to models calibrated on the European (EU) range only. Black dots indicate models calibrated on both EU and North American (NA) ranges. Abs(Δ):μ indicates the average absolute difference between simulated and detected overlaps. The significance of the Wilcoxon signed-rank test, *W*, is shown (ns, non-significant; $^{*}0.05 < P < 0.01$; $^{***}P < 0.001$).

tions, which remove collinearity between variables by finding orthogonal axes, the variable selection procedure of SDMs is sensitive to collinearity. A variable that is not important for the biology of the species, but correlated to one that is, might be given a high weight in the model (e.g. as in the case of microclimatic decoupling of macroclimatic conditions; Scherrer & Körner, 2010). Projection of the model to another area (or continent in the present case) could then be inconsistent with the actual requirements of the species and lead to spurious patterns of detected overlap. In contrast, ordination techniques calibrated on only one study area show a more stable pattern of detected overlap (i.e. monotonic increase, low Δ_{abs}:μ). In general, no SDM method exceeded the performance of the best ordination method.

Based on our results, ordinations seem to be more appropriate than SDMs for investigating niche overlap. However, unlike ordination techniques, SDMs are able to select and rank variables according to their importance in delimiting the niche. SDMs thus could be used to identify variables that are closely related to the processes driving the distribution of the species, while excluding variables that do not discriminate presence and absence. It remains to be tested whether the use

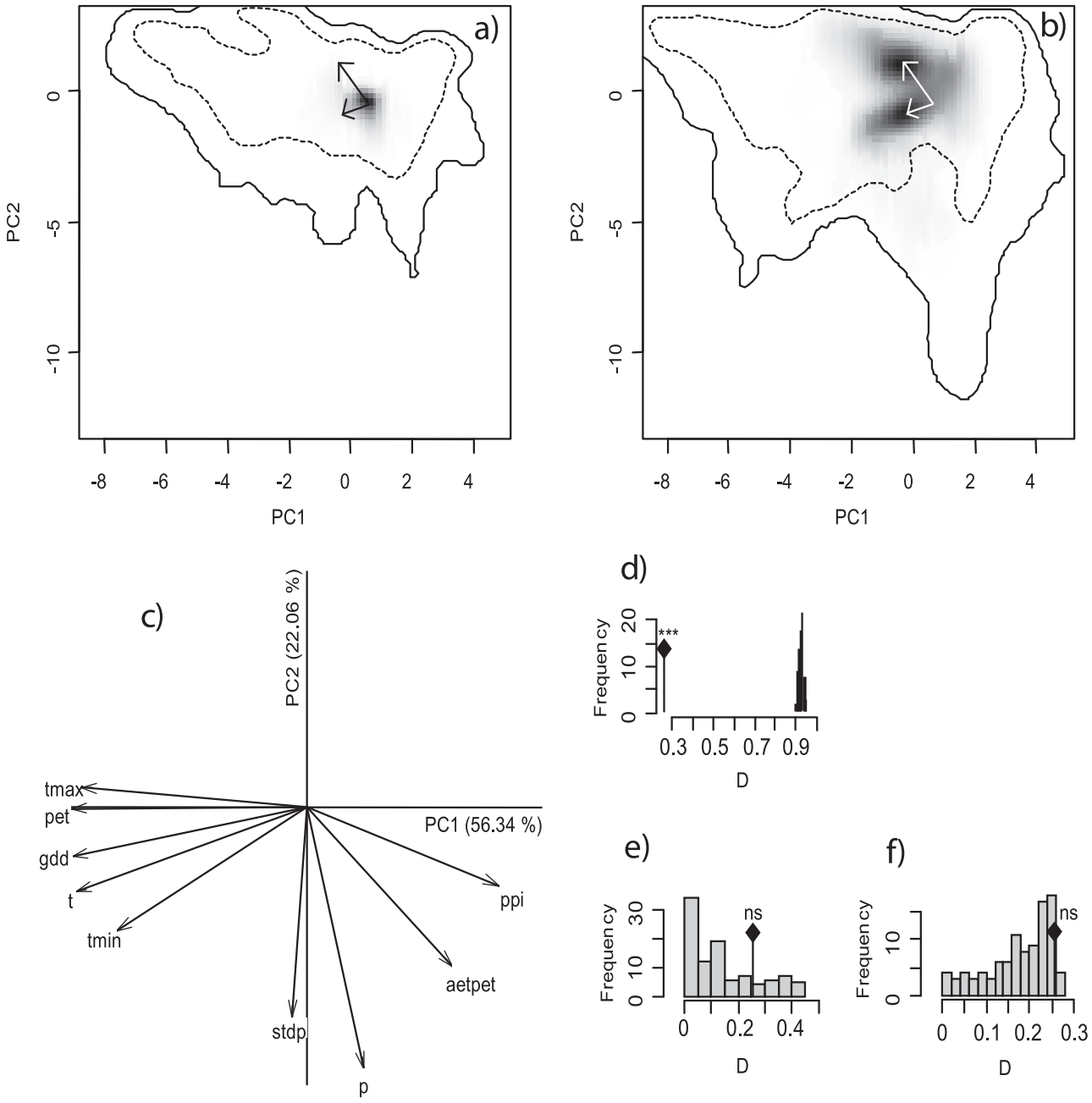


Figure 6 Niche of spotted knapweed in climatic space – example of a principal component analysis (PCA-env). Panels (a) and (b) represent the niche of the species along the two first axes of the PCA in the European native (EU) and North American invaded range (NA), respectively. Grey shading shows the density of the occurrences of the species by cell. The solid and dashed contour lines illustrate, respectively, 100% and 50% of the available (background) environment. The arrows represent how the centre of the niche has changed between EU and NA. (c) The contribution of the climatic variables on the two axes of the PCA and the percentage of inertia explained by the two axes. Histograms (d)–(f) show the observed niche overlap *D* between the two ranges (bars with a diamond) and simulated niche overlaps (grey bars) on which tests of niche equivalency (d), niche similarity of NA to EU (e), and niche similarity of EU to NA (f) are calculated from 100 iterations. The significance of the tests is shown (ns, non-significant; $^{***}P < 0.001$).

of simpler SDM models with more proximal variables (i.e. thus reducing the potential influence of model overfitting and variable collinearity; Guisan & Thuiller, 2005) would improve the accuracy of estimated niche overlap. The best practice is to use variables thought to be crucial (i.e. eco-physiologically meaningful) for the biology of the species (Guisan & Thuiller, 2005). Often, uncertainties surrounding the biology of focal species leave us to select variables relevant to the eco-physiology of the higher taxonomic group to which it belongs (e.g. all vascular plants).

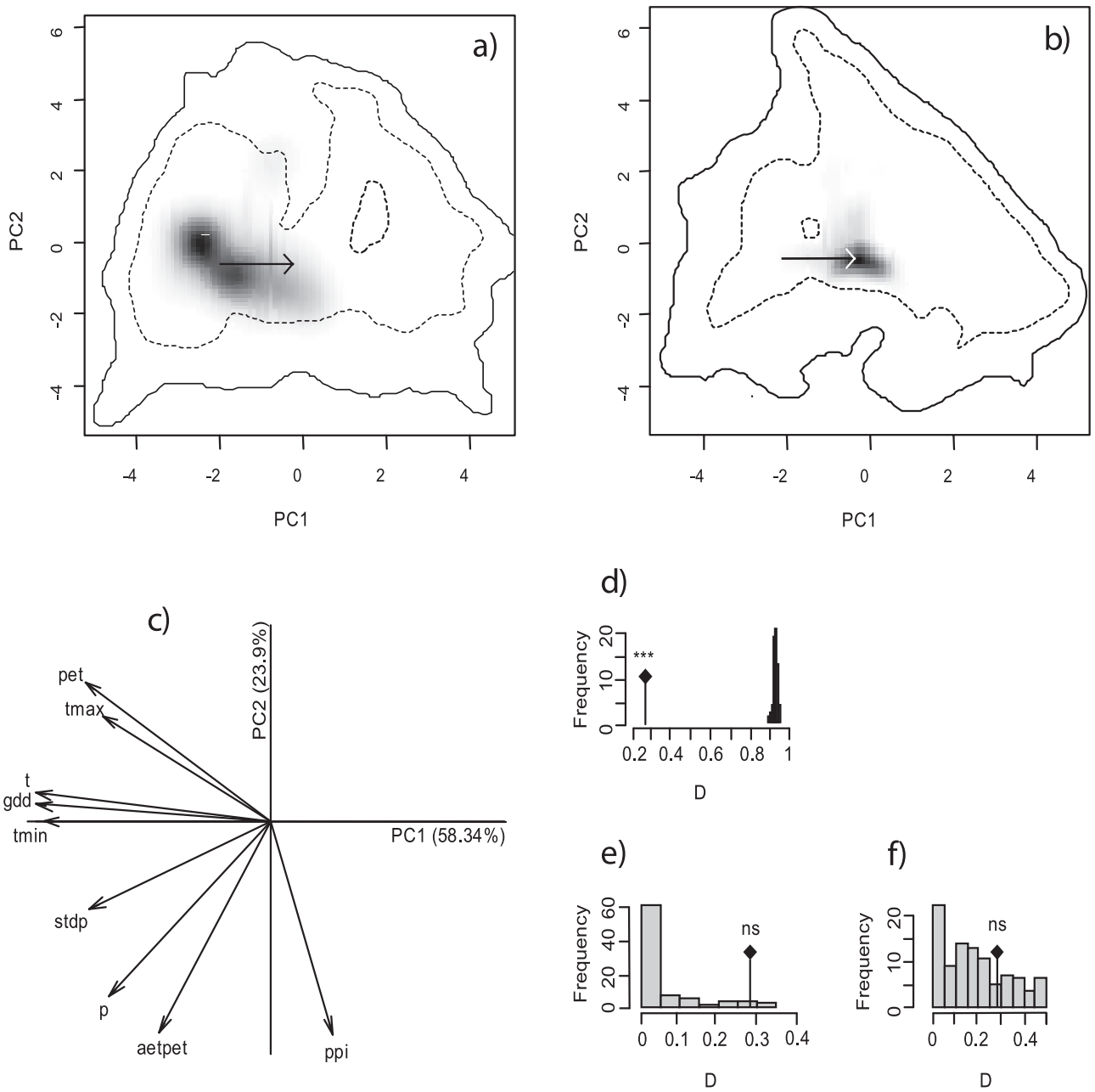


Figure 7 Niche of the red imported fire ant in climatic space – example of a principal component analysis (PCA-env). Panels (a) and b) represent the niche of the species along the two first axes of the PCA in the South American native (SA) and North American invaded range (NA), respectively. Grey shading shows the density of the occurrences of the species by cell. The solid and dashed contour lines illustrate, respectively, 100% and 50% of the available (background) environment. The arrows represent how the centre of the niche has changed between SA and NA. (c) The contribution of the climatic variables on the two axes of the PCA and the percentage of inertia explained by the two axes. Histograms (d)–(f) show the observed niche overlap D between the two ranges (bars with a diamond) and simulated niche overlaps (grey bars) on which tests of niche equivalency (d), niche similarity of NA to SA (e), and niche similarity of SA to NA (f) are calculated from 100 iterations. The significance of the tests is shown (ns, non-significant; $^{***}P < 0.001$).

Differences in overlap detection among ordinations

Of the ordination techniques we considered, PCA-env most accurately quantified the simulated level of niche overlap and did so without substantial bias. Unlike PCA-occ, PCA-env summarizes the entire range of climatic variability found in the study area and it is in this multivariate space that occurrences of the species are then projected. Thus, PCA-env is less prone to artificial maximization of ecologically irrelevant differences between distributions of the species. However, the possibility remains that superior performance of PCA-env might be partly attributable to the fact that our study areas (i.e. Europe and

North America) have relatively similar precipitation and temperature gradients that explain most of the environmental variation. The highest performance of PCA-env is likely in situations where species respond to gradients that also account for most of the environmental variation throughout the study region as a whole (i.e. the maximization of the variation of the environment in the study area also maximizes the variation in the niche of the species). Moreover, if this environmental setting prevails in both study areas, issues regarding changes in the correlation structure of variables may be minimal.

PCA-occ, in contrast, uses environmental values at species occurrences only and selects variables that vary most among occurrences. The resulting principal components are calibrated to discriminate even the slightest differences in the correlation of variables at each occurrence. A variable that differs little among locations where the species occurs, but exhibits substantial variation across the study region, probably represents meaningful ecological constraint. Therefore, depending on the environment of the study region (which PCA-occ does not consider), these variables may have undetected ecological relevance (Calenge *et al*., 2008). If the noise (e.g. climatic variation between regions) is large relative to the signal (i.e. differences in niches between species), the degree of niche overlap could be underestimated (Fig. 3a).

LDA and BETWEEN-occ analyses calibrated using occurrences alone tend to underestimate the simulated level of niche overlap. Both of these methods attempt to discriminate a priori chosen groups along environmental gradients. Therefore, these methods will give a higher weight to variables that discriminate the two niches in terms of average positions. For example, in the case of a perfect overlap between the niches on temperature (t) and precipitation (p) variables, these methods will ignore environmental variables most correlated with t and p, and will instead select variables that discriminate the niches, no matter their ecological relevance. Therefore, these methods will tend to erroneously suggest that niches differ more than they actually do. If such group discriminant analyses show high overlap, there is no difference in the average position of the niches along any variable. However, if they show low overlap, one should be aware of the ecological relevance of the components along which the niche average positions differ.

WITHIN-env was the second most reliable method for quantifying niche overlap. This method aims to first remove differences between the two environments and subsequently focuses on differences between the niches in a common multivariate environmental space. All information that is not shared by the two environments is not retained. This approach is more conservative and therefore may be more robust in analyses where two areas (or times) widely differ regarding some variables. A niche shift detected after removing the effect of the different environments is unlikely to be a statistical artefact and therefore probably represents a true difference or change in the ecology of the species. That said, the superior performance of WITHIN-env in our study is probably related to the manner in which distributions were simulated (equal variance, but different means) and this approach may not perform well if the excluded variables (i.e. the gradients showing largest differences between the two areas) are relevant with respect to niche quantification and, thus, niche overlap between the two distributions. In such cases, only limited conclusions regarding niche differences are possible, although the retained variables may actually be important determinants of the species' niche. In contrast, the WITHIN-occ method (i.e. calibrated on occurrences only) significantly overestimated the simulated degree of overlap. This was expected since the method removes most of the environmental differences found between the two sets of occurrences before comparing the niches. For this reason, we anticipated even greater overestimation of niche overlap.

In the case of ENFA, information is also lost because the two selected axes do not maximize the explained variation. Instead, ENFA constructs the niche using a model with a priori ecological hypotheses that are based on the concepts of marginality and specificity (Hirzel *et al*., 2002). Therefore, ENFA tends to suggest niches are more similar than they actually are.

Despite differences between ordination methods, all were consistent in one aspect. When calibrated on both the EU and NA ranges, the measured niche overlap (filled circles, Fig. 3) was generally lower than the simulated level and also lower than the measured values when calibrated on EU alone (crosses, Fig. 3). When only one range is used in the calibration process, less climatic variation is depicted in the environmental space, thus increasing the overlap between distributions.

Reanalysis of case studies

In the cases of spotted knapweed, *Centaurea stoebe* (Broennimann *et al*., 2007), and the fire ant, *Solenopsis invicta* (Fitzpatrick *et al*., 2007, 2008), niche overlap was originally assessed through the use of a BETWEEN-occ analysis and the calculation of the between-class ratio of inertia that does not correct for environmental availability (spotted knapweed 0.32; fire ant 0.40). Although our framework produced different values of niche overlap with PCA-env (spotted knapweed and fire ant 0.25 and 0.28, respectively; Figs 6 & 7), the conclusions in the original papers do not change. Namely, this reanalysis confirms earlier findings that both spotted knapweed and the fire ant experienced measurable changes in environmental niche occupancy as they invaded North America. The application of our framework to these species results in rejection of the niche equivalency hypothesis. Despite claims to the contrary (e.g. Peterson & Nakazawa, 2008), our analyses confirm that any attempt to predict the niche characteristics from one range to another is inadequate for these species. The results also show that, as would be expected, the invasive niches tend to be more similar to the native niche than random and, thus, niche similarity could not be rejected. In the perspective of niche conservatism we thus conclude that, as invasive species, spotted knapweed and the fire ant did not significantly retain their environmental niche characteristics from their native ranges.

Perspectives

We developed and tested our framework using only one set of study areas comprising all environments present in EU and NA. Virtual entities were created by varying niche positions along environmental gradients but with constant niche breadths. We used this setting, which obviously is a subset of situations encountered in nature, because of computational limitations and to simplify the interpretation of the results. Though we believe that this setting provides robust insights to develop best practices for quantification of niche overlap, other situations should be investigated. To explore differences between ordination and SDM techniques more fully, one would need to simulate species distributions with low to high variance of the environment in the study region as a factor that is crossed with low/high variance of the environmental conditions at species occurrences. We cannot exclude that some modelling technique (i.e. such as MaxEnt, the only SDM method which provided irregular, but non-significantly biased results) could be more robust when differences between environments are important.

The framework we illustrate here measures niche overlap using the metric *D* (Schoener, 1970). Different metrics exist to measure niche overlap (e.g. MacArthur & Levins, 1967; Colwell & Futuyma, 1971; Warren *et al*., 2008) and since we provide a description of the niche in a gridded environmental space, these additional measures or metrics could be easily implemented. However we feel that the metric *D* is the easiest to interpret. This measure indicates an overall match between two niches over the whole climatic space and determines whether we can infer the niche characteristics of one species (subspecies, population) from the other. We argue that SDMs can be reasonably projected outside the calibration area only if the niche overlap is high ($D \approx 1$) and if the test of niche equivalency could not be rejected.

The metric *D* (as most overlap metrics) does not indicate the directionality or type of niche difference and alone cannot tell us whether the niche has expanded, shrunk or remained unchanged. In a similar vein, because *D* is symmetrical, the amount of overlap is the same for both entities being compared, even though it is unlikely that the niches of two entities are of the same size. Moreover, *D* provides no quantitative indication concerning the position and the breadth of the niches (but does provide a visual indication). These additional measures of the directionality of niche change could be easily implemented in our framework in the future.

CONCLUSIONS

How the environmental niches of taxa change across space and time is fundamental to our understanding of many issues in ecology and evolution. We anticipate that such knowledge will have practical importance as ecologists are increasingly asked to forecast biological invasions, changes in species distributions under climatic change or extinction risks. To date, our ability to rigorously investigate intra- or inter-specific niche overlap has been plagued by methodological limitations coupled with a lack of clarity in the hypotheses being tested. The result has been ambiguity in interpretation and inability to decipher biological signals from statistical artefacts. The framework we present allows niche quantification through ordination and SDM techniques while taking into account the availability of environments in the study area. As in Warren *et al*. (2008), our framework allows statistical tests of niche hypotheses (i.e. niche similarity and equivalency), but under our framework these tests are performed directly in environmental space, thereby allowing correction of bias associated with geographical dimension. Our comparative analysis of virtual entities with known amounts of niche overlap shows that such ordination techniques quantify niche overlap more accurately than SDMs. However, we show that the choice of technique, depending on the structure of the data and the hypotheses to test, remains critical for an accurate assessment of niche overlap. Focusing on rates of change of species niches and a search for consistent patterns of niche lability and/or stability across many taxa will most readily complement the synthesis of ecological and evolutionary analyses already firmly under way.

ACKNOWLEDGEMENTS

O.B. was funded by the National Centre of Competence in Research (NCCR) Plant Survival, a research program of the Swiss National Science Foundation. M.C.F. was partially supported by funding from UMCES. P.B.P, A.G., N.G.Y., N.E.Z, W.T. were supported by the 6th European Framework Programme grants GOCE-CT-2007?036866 'ECOCHANGE'. P.B.P. acknowledges support from Swiss National Fond grants CRSII3_125240 'SPEED' and 31003A_122433 'ENNIS', and the 7th European Framework ENV-CT-2009–226544 'MOTIVE' during preparation of the manuscript. We thank R. Pearson, J. Soberón and two anonymous referees for valuable comments and insights on a previous version of the manuscript.

REFERENCES

Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.

Blair, W.F. (1940) Home ranges and populations of the jumping mouse. *American Midland Naturalist*, **23**, 244–250.

Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5–32.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, **4**, 585–589.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701–709.

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

Chase, J.M. & Leibold, M.A. (2003) *Ecological niche: linking classical and contemporary approaches*. The University of Chicago Press, Chicago.

Colwell, R.K. & Futuyma, D.J. (1971) Measurement of niche breadth and overlap. *Ecology*, **52**, 567–576.

Colwell, R.K. & Rangel, T.F. (2009) Hutchinson's duality: the once and future niche. *Proceedings of the National Academy of Sciences USA*, **106**, 19651–19658.

Devictor, V., Clavel, J., Julliard, R., Lavergne, S., Mouillot, D., Thuiller, W., Venail, P., Villéger, S. & Mouquet, N. (2010) Defining and measuring ecological specialization. *Journal of Applied Ecology*, **47**, 15–25.

Dolédec, S. & Chessel, D. (1987) Seasonal successions and spatial variables in freshwater environments. 1. Description of a complete 2-way layout by projection of variables. *Acta Oecologica–Oecologia Generalis*, **8**, 403–426.

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

Etterson, J.R. & Shaw, R.G. (2001) Constraint to adaptive evolution in response to global warming. *Science*, **294**, 151–154.

Evans, M.E.K., Smith, S.A., Flynn, R.S. & Donoghue, M.J. (2009) Climate, niche evolution, and diversification of the 'bird-cage' evening primroses (*Oenothera*, sections *Anogra* and *Kleinia*). *The American Naturalist*, **173**, 225–240.

Ferrier, S., Manion, G., Elith, J. & Richardson, K. (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions*, **13**, 252–264.

Fisher, R.A. (1936) The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, **7**, 179–188.

Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, **16**, 24–33.

Fitzpatrick, M.C., Dunn, R.R. & Sanders, N.J. (2008) Data sets matter, but so do evolution and ecology. *Global Ecology and Biogeography*, **17**, 562–565.

Friedman, J.H. (2001) Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, **29**, 1189–1232.

Gaston, K.J. (2003) *The structure and dynamics of geographic ranges*. Oxford University Press, Oxford.

Gotelli, N.J., Graves, G.R. & Rahbek, C. (2010) Macroecological signals of species interactions in the Danish avifauna. *Proceedings of the National Academy of Sciences of the USA*, **107**, 5030–5035.

Gower, J.C. (1966) Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, **53**, 325–328.

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, **19**, 497–503.

Grinnell, J. (1917) The niche-relationships of the California thrasher. *The Auk*, **34**, 427–433.

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

Hengl, T., Sierdsema, H., Radović, A. & Dilo, A. (2009) Spatial prediction of species' distributions from occurrence-only records: combining point pattern analysis, ENFA and regression-kriging. *Ecological Modelling*, **220**, 3499–3511.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

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.

Hof, C., Rahbek, C. & Araújo, M.B. (2010) Phylogenetic signals in the climatic niches of the world's amphibians. *Ecography*, **33**, 242–250.

Horn, H.S. (1966) Measurement of overlap in comparative ecological studies. *The American Naturalist*, **100**, 419–424.

Hutchinson, G.E. (1957) Population studies – animal ecology and demography – concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.

Jongman, R.H.G., ter Braak, C.J.F. & Van Tongeren, O.F.R. (1995) *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge.

Jump, A.S. & Peñuelas, J. (2005) Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters*, **8**, 1010–1020.

Losos, J.B., Leal, M., Glor, R.E., de Queiroz, K., Hertz, P.E., Schettino, L.R., Lara, A.C., Jackman, T.R. & Larson, A. (2003) Niche lability in the evolution of a Caribbean lizard community. *Nature*, **424**, 542–545.

Lütolf, M., Bolliger, J., Kienast, F. & Guisan, A. (2009) Scenario-based assessment of future land use change on butterfly species distributions. *Biodiversity and Conservation*, **18**, 1329–1347.

MacArthur, R. & Levins, R. (1967) Limiting similarity, convergence, and divergence of coexisting species. *The American Naturalist*, **101**, 377–385.

McCullagh, P. & Nelder, J. (1989) *Generalized linear models*, 2nd edn. Chapman and Hall/CRC, Boca Raton, FL.

May, R.M. & Arthur, R.H.M. (1972) Niche overlap as a function of environmental variability. *Proceedings of the National Academy of Sciences USA*, **69**, 1109–1113.

Medley, K.A. (2010) Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus* Skuse (Culicidae), revealed by reciprocal distribution models. *Global Ecology and Biogeography*, **19**, 122–133.

New, M., Lister, D., Hulme, M. & Makin, I. (2002) A high-resolution data set of surface climate over global land areas. *Climate Research*, **21**, 1–25.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008a) Niche dynamics in space and time. *Trends in Ecology and Evolution*, **23**, 149–158.

Pearman, P.B., Randin, C.F., Broennimann, O., Vittoz, P., van der Knaap, W.O., Engler, R., Le Lay, G., Zimmermann, N.E. & Guisan, A. (2008b) Prediction of plant species distributions across six millennia. *Ecology Letters*, **11**, 357–369.

Pearson, K. (1901) On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, **2**, 559–572.

Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with

Solenopsis invicta and *Solenopsis richteri*. *Global Ecology and Biogeography*, **17**, 135–144.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Pianka, E.R. (1980) Guild structure in desert lizards. *Oikos*, **35**, 194–201.

R Development Core Team (2010) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Scherrer, D. & Körner, C. (2010) Infra-red thermometry of alpine landscapes challenges climatic warming projections. *Global Change Biology*, **16**, 2602–2613.

Schoener, T.W. (1970) Nonsynchronous spatial overlap of lizards in patchy habitats. *Ecology*, **51**, 408–418.

Silverman, B.W. (1986) *Density estimation for statistics and data analysis*. Chapman and Hall, London.

Silvertown, J., Dodd, M. & Gowing, D. (2001) Phylogeny and the niche structure of meadow plant communities. *Journal of Ecology*, **89**, 428–435.

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

Steiner, F.M., Schlick-Steiner, B.C., VanDerWal, J., Reuther, K.D., Christian, E., Stauffer, C., Suarez, A.V., Williams, S.E. & Crozier, R.H. (2008) Combined modelling of distribution and niche in invasion biology: a case study of two invasive *Tetramorium* ant species. *Diversity and Distributions*, **14**, 538–545.

Sunday, J.M., Bates, A.E. & Dulvy, N.K. (2011) Global analysis of thermal tolerance and latitude in ectotherms. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 1823–1830.

Thuiller, W., Lavorel, S. & Araújo, M.B. (2005a) Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology and Biogeography*, **14**, 347–357.

Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005b) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234–2250.

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

Varela, S., Lobo, J.M., Rodríguez, J. & Batra, P. (2010) Were the Late Pleistocene climatic changes responsible for the disappearance of the European spotted hyena populations? Hindcasting a species geographic distribution across time. *Quaternary Science Reviews*, **29**, 2027–2035.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868–2883.

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

Wintle, B.A., McCarthy, M.A., Volinsky, C.T. & Kavanagh, R.P. (2003) The use of Bayesian model averaging to better represent uncertainty in ecological models. *Conservation Biology*, **17**, 1579–1590.

Worton, B.J. (1989) Kernel methods for estimating the utilization distribution in home-range studies. *Ecology*, **70**, 164–168.

Yesson, C. & Culham, A. (2006) Phyloclimatic modeling: combining phylogenetics and bioclimatic modeling. *Systematic Biology*, **55**, 785–802.

Zimmermann, N.E., Yoccoz, N.G., Edwards, T.C., Meier, E.S., Thuiller, W., Guisan, A., Schmatz, D.R. & Pearman, P.B. (2009) Climatic extremes improve predictions of spatial patterns of tree species. *Proceedings of the National Academy of Sciences USA*, **106**, 19723–19728.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Zip file of scripts.
Appendix S2 Creating virtual entities.
Figure S1 Detected niche overlap as a function of the resolution of the gridded environmental space.
Figure S2 Sensitivity analysis of simulated versus measured niche overlap.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

This work originates from a workshop on 'Progress in predictive species distribution modeling', held in 2008 in Riederalp, Switzerland. O.B., M.C.F., P.B.P. and A.G. conceived the ideas; O.B. wrote the scripts and performed the simulations and analyses; O.B., M.C.F. and P.B.P. led the writing; O.B. and M.C.F. provided data for the two invasive species used for illustrating the approach; B.P. tested the script on further species data not presented here; N.G.Y. provided statistical support; A.G., C.H.G., B.P., L.P., N.G.Y., W.T., M.J.F., C.R. and N.E.Z. suggested important corrections to the manuscript.

Editor: José Alexandre F. Diniz-Filho

Journal of Biogeography (J. Biogeogr.) (2014)

ORIGINAL ARTICLE

Contrasting spatio-temporal climatic niche dynamics during the eastern and western invasions of spotted knapweed in North America

Olivier Broennimann[1]*, Patrik Mráz[2,3], Blaise Petitpierre[1], Antoine Guisan[1,4†] and Heinz Müller-Schärer[2†]

[1]Spatial Ecology Laboratory, Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland, [2]Unit of Ecology and Evolution, Department of Biology, University of Fribourg, Fribourg, Switzerland, [3]Department of Botany, Charles University in Prague, Prague, Czech Republic, [4]Institute of Earth Surface Dynamics, University of Lausanne, Lausanne, Switzerland

*Correspondence: O. Broennimann, Spatial Ecology Laboratory, Department of Ecology and Evolution, University of Lausanne, Biophore Building, CH-1015, Lausanne, Switzerland.
E-mail: olivier.broennimann@unil.ch
†Shared last authorship.

ABSTRACT

Aim The spotted knapweed (*Centaurea stoebe*), a plant native to south-east and central Europe, is highly invasive in North America. We investigated the spatio-temporal climatic niche dynamics of the spotted knapweed in North America along two putative eastern and western invasion routes. We then considered the patterns observed in the light of historical, ecological and evolutionary factors.

Location Europe and North America.

Methods The niche characteristics of the east and west invasive populations of spotted knapweed in North America were determined from documented occurrences over 120 consecutive years (1890–2010). For this investigation, the 2.5 and 97.5 percentiles of values along temperature and precipitation gradients, as given by the two first axes of a principal components analysis (PCA), were calculated. We additionally measured the climatic dissimilarity between invaded sites and the native niche using a multivariate environmental similarity surface (MESS) analysis.

Results Along both invasion routes, the species established in regions with climatic conditions that were similar to those in the native niche. An initial spread in ruderal habitats always preceded spread in (semi-)natural habitats. In the east, the niche gradually increased over time until it reached limits similar to the native niche. Conversely, in the west the niche abruptly expanded after an extended time lag into climates not occupied in the native range; only the native cold niche limit was conserved.

Main conclusions Our study reveals that different niche dynamics have taken place during the eastern and western invasions. This pattern indicates different combinations of historical, ecological and evolutionary factors in the two ranges. We hypothesize that the lack of a well-developed transportation network in the west at the time of the introduction of spotted knapweed confined the species to a geographically and climatically isolated region. The invasion of dry rangelands may have been favoured during the agricultural transition in the 1930s by release from natural enemies, local adaptation and less competitive vegetation, but further experimental and molecular studies are needed to explain these contrasting niche patterns fully. Our study illustrates the need and benefit of applying large-scale, temporally explicit approaches to understanding biological invasions.

Keywords

***Centaurea stoebe*, herbarium records, human disturbances, invasion routes, niche conservatism, niche limits, North America, plant invasions, spotted knapweed, temporal data.**

© 2014 John Wiley & Sons Ltd

Journal of Biogeography

INTRODUCTION

Despite the rapidly accumulating literature on biological invasions, including ever-emerging new hypotheses and meta-analyses, ecologists still lack a detailed understanding of why specific plant invasions occur, making it difficult to predict which species may become invasive and where (Dietz & Edwards, 2006; Gurevitch *et al.*, 2011). Many different mechanisms promoting invasion have been proposed and tested (Dietz & Edwards, 2006), with contradictory findings. Inconsistencies in these findings may be because investigations of mechanisms promoting invasions have been performed indiscriminately at various invasion phases (e.g. lag phase, primary invasion and secondary invasion), locations (e.g. core area of the native range, introduction area and invasion front), habitat types and environmental conditions, each of which is likely to show different invasion dynamics. Accounting for these changes over time holds the promise of a better spatial and temporal understanding of invasions (Dietz & Edwards, 2006), but to our knowledge has not been tried so far. For this to be accomplished, integrative studies are needed that assess key ecological factors throughout the course of invasions.

One key determinant of invasions is the pre-adaptation of species to the environment in the new range, as determined by their climatic niche (Maron *et al.*, 2004; Treier *et al.*, 2009; Di Febbraro *et al.*, 2013). Since the development of niche-based species distribution modelling (SDM; Guisan & Thuiller, 2005; also called ecological niche models, ENM; Peterson *et al.*, 2011), our understanding of abiotic components driving invasions through space has improved considerably (e.g. Gallien *et al.*, 2010). An increasing number of studies have used SDMs (e.g. Broennimann *et al.*, 2007; Fitzpatrick *et al.*, 2007; Rödder *et al.*, 2009; Medley, 2010; Petitpierre *et al.*, 2012) or related tools (e.g. Broennimann *et al.*, 2012) to investigate the niche of invader species and assess whether it is conserved between ranges. However, no investigation has been performed on the mode and tempo of niche changes along invasion routes.

A good model system for this investigation is spotted knapweed (*Centaurea stoebe* L.), a plant native to south-east and central Europe and highly invasive in North America (Sheley *et al.*, 1998). The first records of the species in North America indicate that the species was introduced in the USA near Westford, Massachusetts, in 1884 (Invasive Plant Atlas of New England; http://www.eddmaps.org/ipane; accessed on 13 December 2010) and at the border of the USA and Canada in Victoria, British Columbia, in 1893 (Roche *et al.*, 1986); further records of the plant in the next decade are from the surrounding areas. A recent study has reconstructed the most parsimonious invasion routes using a minimum cost arborescence algorithm (MCA; Hordijk & Broennimann, 2012) and revealed two separate invasions in North America, one starting from the east and the other from the west of the USA and Canada. The availability of a large amount of ecological, distributional and cytological data and the presence of two independent invasions make spotted knapweed a useful unplanned, natural experiment to investigate how ecological and evolutionary factors might interact to result in the observed niche dynamics during biological invasions (Sax *et al.*, 2007).

Recent SDM studies have shown that the initial introduction of spotted knapweed in North America took place within climates similar to those in the native niche (Broennimann *et al.*, 2007; Broennimann & Guisan, 2008) but that the extent of the invasion could not be predicted when projecting models calibrated with the native range. This indicated that the species was able to invade areas with climatic conditions that differed from the native range, in particular into drier climates (Broennimann *et al.*, 2007; Treier *et al.*, 2009). The niche shift of spotted knapweed has been shown to be an exception among Holarctic plant invaders (Petitpierre *et al.*, 2012); it is therefore a particularly suitable model species for integrating a test for niche changes with a comprehensive and multidisciplinary approach that considers invasion routes and detailed knowledge of the species' biology. This integration is required to unravel the underlying evolutionary and ecological processes and to assess their importance at each phase of the plant invasion process.

Here, we present a novel approach to quantify niche changes through time along the two invasion routes and we interpret these findings in the light of known historical processes and experimental studies. We tested this framework using an extensive dataset of occurrences of spotted knapweed that covers the complete distribution of the species in Europe and North America, and we included comprehensive climatic data for these two ranges. For each species record, we gathered distributional, ecological and cytological information that was expected to correlate with mechanisms influencing niche changes at different phases of the invasion process. We also recorded whether the sample was collected in (semi-)natural or ruderal habitats (see Materials and Methods), because in its native range the tetraploid cytotype of spotted knapweed (i.e. the cytotype that colonized North America; Treier *et al.*, 2009) preferentially occurs in drier and more open microhabitats created by human-induced disturbances, while its diploid progenitor occupies (semi-) natural habitats with denser vegetation (Mráz *et al.*, 2012). These disturbances may have increased the likelihood that the plant was transported to a new range and facilitated the successful establishment of new populations and subsequent invasions (anthropogenically induced adaptation to invade theory; Hufbauer *et al.*, 2012).

Our main objective was to provide the first analysis of niche changes in space and time of a major plant invader, spotted knapweed, across its successive invasion stages in North America. More specifically, we asked the following questions. (1) Did climatic niche limits change through time during the course of the invasion and did they differ between the two invasion routes? (2) Did habitat types colonized by

© 2014 John Wiley & Sons Ltd

the species change through time and did populations in (semi-)natural and ruderal habitats show different niche limits, and were these patterns different for the two invasion routes? (3) What can we learn about the observed patterns of niche dynamics in the light of historical, ecological and evolutionary processes evidenced in previous studies on this species? We present the results in relation to questions (1) and (2), and then discuss the results in the light of question (3).

MATERIALS AND METHODS

The species

Spotted knapweed occurs in its native range as a diploid and tetraploid cytotype but only the latter has been recorded in its invaded range (Treier *et al.*, 2009; Mráz *et al.*, 2011; Fig. 1). Such a pronounced shift in cytotype frequency has been explained either by stochastic founder events or by a superior establishment, colonization and persistence ability of tetraploids compared with diploids, since tetraploids are predominantly short-lived perennial polycarpic plants, while diploids are annual-biennial monocarps (Müller-Schärer *et al.*, 2004; Henery *et al.*, 2010; Hahn *et al.*, 2012).

Occurrence data

We began by updating the database for *Centaurea stoebe* (Asteraceae) in Europe and North America previously used in Broennimann *et al.* (2007) and Broennimann & Guisan (2008), mainly by expanding it to include data from Russia and eastern Europe as well as eastern North America. We meticulously searched local and regional herbaria and databases available online (see Appendix S1 in Supporting Information), revised herbarium vouchers in selected European herbaria (Appendix S1), included all available cytogeographical data (P. Mráz *et al.*, unpublished data), and revised distributional data published by Ochsmann (2000). We only recorded occurrences for which ploidy level, habitat description and date of collection were available. When cytological data were not directly available for a specimen, we used scans of herbarium vouchers to ascribe the ploidy level based on morphological characteristics (Mráz *et al.*, 2011). Using these data, we created separate datasets for each of the four geo-cytotypes, the European diploid i.e. *C. stoebe* s. str., the European tetraploid, i.e. *C. stoebe* s. lat., the North American tetraploid from the east coast and the North American tetraploid from the west coast. North American populations were attributed to the east or west coast invasion according to the reconstruction of the invasion routes (Hordijk & Broennimann, 2012; see also Fig. 2). These datasets provide the most exhaustive historical and distributional information on spotted knapweed to date, with 3631 occurrences recorded between 1831 and 2010, with good coverage of the areas in which the species has been reported (Fig. 1).

Natural and ruderal habitats

The habitat information was classified according to the European classification system of habitats (EUNIS, 2008). Information gathered from herbaria, online databases or from our own sampling were carefully analysed, first using keywords from the habitat descriptions in EUNIS, then using Google Earth in cases of ambiguity or lack of information. The habitats belonging to the categories of diluvial sediments (C), natural and semi-natural grasslands (E) and natural rocky outcrops (H) were classified as (semi-)natural, whereas habitats belonging to agricultural habitats: fields, vineyards (I) and artificial and industrial habitats: transport networks, extractive industrial sites (J) were interpreted as ruderal habitats.

Climatic data

We used eight bioclimatic variables at 10′ resolution derived from raw climatic data from the CRU CL 2.0 dataset (New *et al.*, 2002): ratio of actual and potential evapotranspiration (*aet/pet*); number of growing degree-days above 5 °C (*gdd*); annual precipitation (*p*); potential evapotranspiration (*pet*); number of months with drought (*ppi*); seasonality of precipitation (*stdp*); annual mean temperature (*t*); annual maximum temperature (t_{max}); and annual minimum temperature (t_{min}). The CRU CL 2.0 data covered 1961–1990, which was not the entire timeframe examined in the study, but it was considered that this bias would be negligible with regard to the amplitude of the climatic gradients covered by the species (e.g. *c.* 15 °C for annual mean temperature). The procedures for calculating *aet*/*pet*, *pet* and *gdd* from the raw CRU CL 2.0 data are detailed in Thuiller *et al.* (2005). The study area (the grey areas in Fig. 1) was selected to include all the biomes occupied by the species in Europe (EU) and North America (NA).

Statistical analyses

We performed a principal components analysis (PCA) on all the climatic variables covering all the sites in the study area using the ade4 library in R (Dray & Dufour, 2007). This produced an orthonormal system of principal components that maximized the environmental variation present in the study area. This method has been shown to be appropriate for quantifying niche overlap (Broennimann *et al.*, 2012). The first two axes accounted for 78.4% of the inertia of the PCA and corresponded broadly to a gradient of temperature and a gradient of humidity (see Appendix S2). Occurrences were then projected along the axes of this orthogonal environmental space. For each year and for each geo-cytotype, we considered all occurrences collected until that year and calculated the 2.5 and 97.5 percentiles of their scores along the two first axes as the lower and upper niche limits. This ensured that niche limits were defined by 95% of occurrences while the remaining 5% were considered to be outliers. Note that, with this

© 2014 John Wiley & Sons Ltd

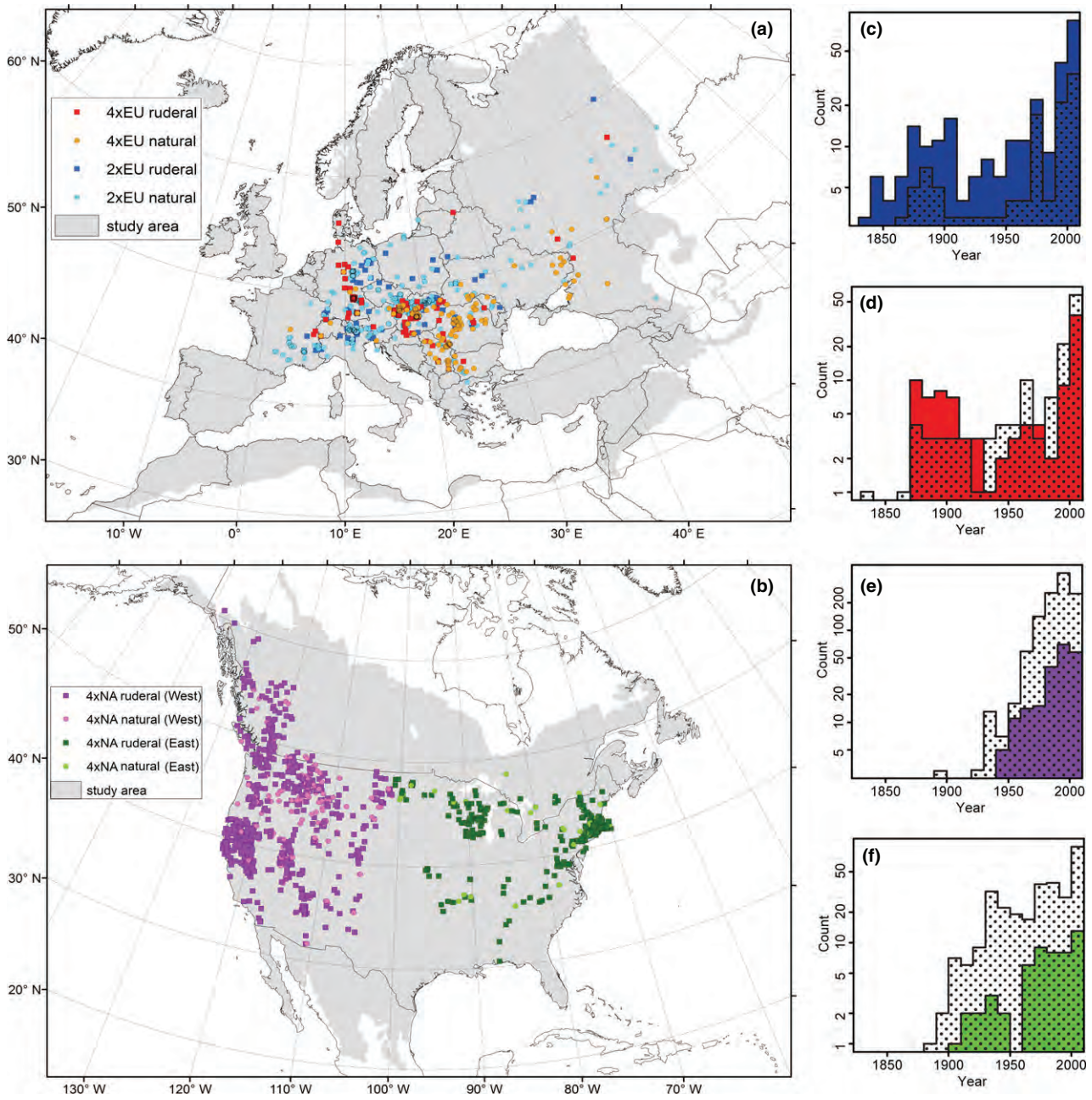


Figure 1 Occurrences of spotted knapweed (*Centaurea stoebe*) in its native European and introduced North American ranges. Spatial distributions of occurrences where dates of collection, ploidy level and description of habitat are known are shown for the native (a) and introduced (b) ranges. Populations found in (semi-)natural habitats are indicated with circles; those in ruderal habitats are indicated with squares. The number of occurrences in ruderal (dotted areas) and natural (solid areas) habitats per time slices of 10 years are shown for the European diploid geo-cytotype (2×EU) (c), the European tetraploid (4×EU) (d) and the North American tetraploid (4×NA) from the west coast (e) and east coast (f). Colours given in (a) and (b) correspond to those in (c–f).

procedure, we assumed that the species did not disappear from a site once colonized. This is a reasonable assumption for this highly invasive plant.

The niche limits were calculated only for years with more than five records. For each geo-cytotype, we performed non-parametric Wilcoxon tests to assess whether the niche limits were significantly different between populations growing in (semi-)natural and ruderal habitats. For each geo-cytotype, we broadly defined a lag phase (i.e. the period of time between the first introduction and five populations), a spread phase (i.e. the period of time after the lag phase with only ruderal populations) and an expansion phase (i.e. the period of time after the initial phase with both ruderal and (semi-) natural populations) (Fig. 3).

Furthermore, we measured the climatic dissimilarity of each site of the study area compared with the native niche.

© 2014 John Wiley & Sons Ltd

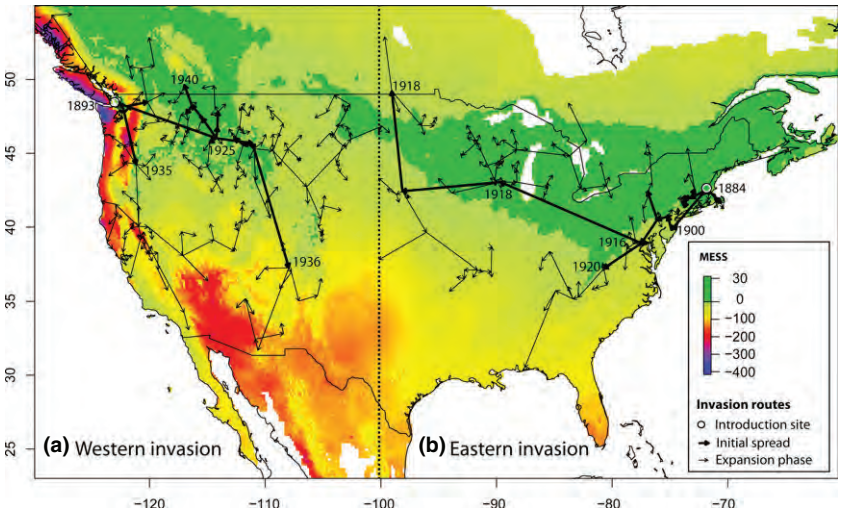


Figure 2 Pattern of spread and multivariate environmental similarity surface (MESS) analysis. A reconstruction of invasion routes originating from the western (a) and eastern (b) introduction sites is shown with arrows (from Hordijk & Broennimann, 2012). Thick and thin arrows correspond to the initial spread and expansion phases, respectively, as shown in Fig. 3. Dark green areas indicate sites with climatic conditions similar to those in the native niche (positive MESS values; note that both introduction sites are in green). The light green–blue gradient indicates the degree of dissimilarity with the climate of the native distribution (negative MESS values).

To do so, we computed multivariate environmental similarity surfaces (MESS; Elith *et al.*, 2010) using the library DISMO in R. The MESS analysis provided a representation of the similarity of grid cells (i.e. with the same resolution and extent as the climatic data) occupied by the species in the invaded range compared with the grid cells occupied by the species in the native range (thus defining the native niche), with respect to the set of environmental variables. A grid cell with a positive value indicated that it fell within the range of environmental values of the native niche, while a grid cell with a negative value indicated that at least one variable had a value that was outside the range of environmental values of the native niche.

RESULTS

Introductions to the new ranges and lag time

Both presumed introduction locations showed climatic conditions similar to those in the native niche (Fig. 3) and have been predicted to be suitable for the species by previous distribution models calibrated with native occurrences (Broennimann *et al.*, 2007; Broennimann & Guisan, 2008). The introduction in the west took place in slightly wetter conditions than in the native niche, according to the second axis of the PCA (Fig. 3d). Overall, however, our results, together with a number of other studies (Broennimann *et al.*, 2007; Fitzpatrick *et al.*, 2007; Rödder *et al.*, 2009; Medley, 2010), provided evidence that sites of successful introductions matched conditions found within the native niche of the species. The MESS analysis also confirmed this because introduction sites showed positive values (Fig. 2).

Spotted knapweed did not spread immediately after being introduced into a new range. The distribution of the species remained confined to a few populations for about 20 years in the east, and about 40 years in the west (Fig. 3). In the west, the species remained restricted to the Victoria coastal area (British Columbia, Canada; Roche *et al.*, 1986), an isolated patch of wetter pacific climate surrounded by drier regions (Fig. 2a). This lag time is relatively short compared with most plants that have invaded temperate climates (i.e. Crooks, 2005; Daehler, 2009).

Initial spread in the introduced range

The initial spread and niche expansions of spotted knapweed in both the eastern and western invaded ranges were achieved almost exclusively in ruderal habitats (Fig. 3). Colonization of natural and semi-natural habitats started only after *c.* 30 years in the east (*c.* 10 years after the beginning of its spread in ruderal habitats) and after *c.* 60 years in the west (*c.* 20 years after the beginning of its spread in ruderal habitats) (Fig. 3). Most of the current eastern range of the species was already covered by the end of this initial spread phase, around 1920. In contrast, the initial invasion was much slower in the west, and only the north-western part of the range was occupied by the end of the 1950s, notably in Montana, USA, where several successive spread events occurred in the 1920s and 1930s (Fig. 2). Interestingly, the slow velocity of the invasion in the west matched the pattern of climatic distance from the native niche, reflecting the fact that it probably took more time for the species to establish in habitats that were quite different compared with the ones present at the introduction site (Fig. 2a). During this initial spread phase, niche limits in both ranges remained similar to

© 2014 John Wiley & Sons Ltd

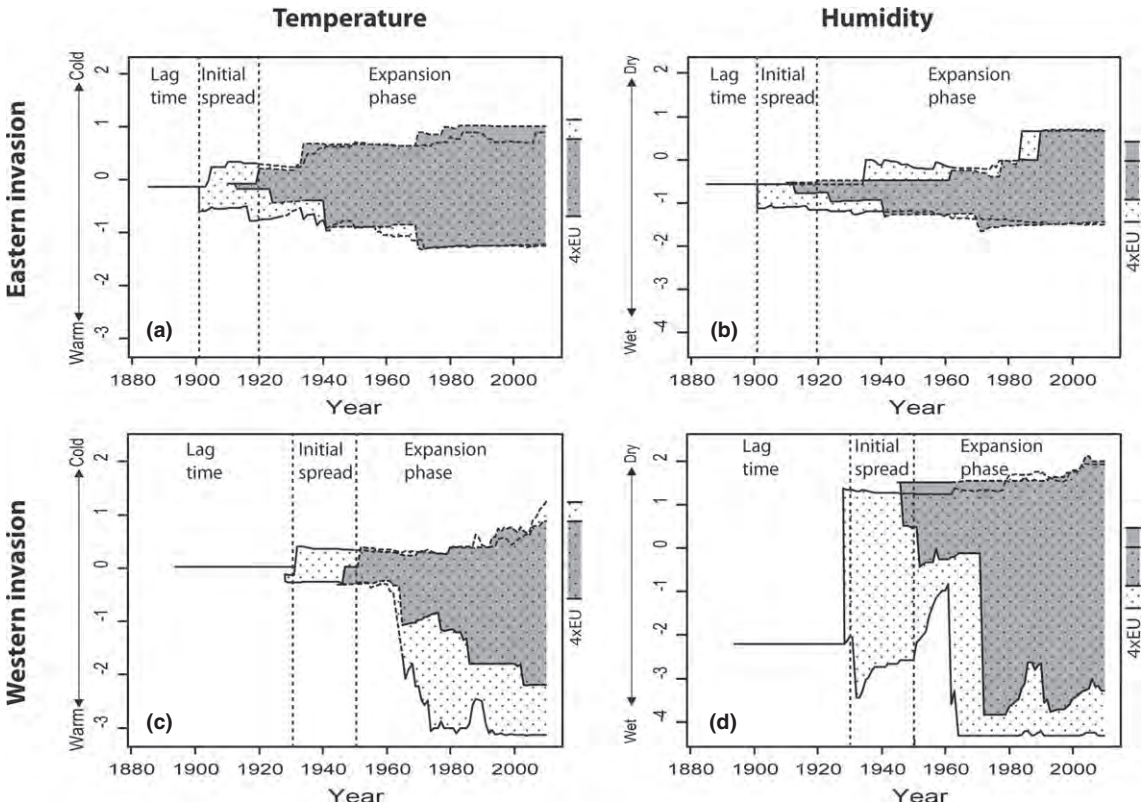


Figure 3 Dynamics of niche limits over time**.** The left panels show the change in the realized niche along a temperature gradient [first principal components analysis (PCA) axis, 56.3% of explained variance]; the right panels show the change in the niche along a humidity gradient (second PCA axis, 22.1% of explained variance). (a, b) The niche limits over time for the eastern North American tetrapoid geo-cytotype (4×NA) populations; (c, d) the niche limits over time for the western 4×NA populations. Lines represent the 2.5% upper and lower quantiles. Solid and dashed lines indicate significant and non-significant differences, respectively, in niche limits between populations growing in natural (grey areas) and ruderal (dotted areas) habitats. Horizontal tick marks on the right side of the plots indicate the niche limits of the European tetraploid geo-cytotype (4×EU) for visual comparison.

those in the native niche, except for the humidity gradient in the western invasion, where drier habitats were suddenly colonized (i.e. mostly in Montana).

Main expansion and niche dynamics in the introduced range

Distinct niche limit patterns could be seen along the eastern and western routes. In the east, populations occupying (semi-)natural and ruderal habitats (Fig. 3a,b) gradually spread to colonize cold, dry and wet niche limits very similar to those of the European tetraploids, with only the warm niche limit being slightly altered compared with the native niche (Fig. 3a). This reflected the fact that new populations were able to establish in regions neighbouring those already colonized during the early spread phase or new regions with climatic conditions relatively similar to those in the introduction site (Fig. 2b). In contrast, the realized niche abruptly expanded in the western invasion, displacing the niche limits along the humidity gradient towards both ends, i.e. drier and wetter conditions (Fig. 3d), and then also gradually along the temperature gradient towards significantly warmer conditions; the cold limit remained similar to that in the native niche (Fig. 3c). These niche limit expansions reflected the observed geographical spread after the 1950s towards the Dakotas (east USA), Colorado (south-east USA) and California (south USA), all of which provide different climatic conditions (Fig. 2a). During the last 30 years, the niche limits have not changed further.

DISCUSSION

Evidence of species' niche shifts during biological invasions are increasingly being reported (e.g. Broennimann *et al.*, 2007; Fitzpatrick *et al.*, 2007; Rödder *et al.*, 2009; Medley, 2010; Petitpierre *et al.*, 2012). However, Petitpierre *et al.* (2012) have shown that climatic niche shifts are actually rare among Holarctic plant invaders, but identified spotted knapweed as a remarkable outlier, confirming earlier findings by Broennimann *et al.* (2007). Unlike previous studies, here we have used novel temporal data and niche analyses to (1) reveal large expansions through time of the realized niche of spotted knapweed in its invaded range (Fig. 3), and (2) identify the time and mode of these niche expansions. This

© 2014 John Wiley & Sons Ltd

temporally and spatially explicit dataset allowed us to show that different niche dynamics took place along distinct invasion routes in eastern and western North America. Below, we discuss the potential mechanisms and possible processes underlying the observed patterns of niche dynamics in the light of the historical, ecological and evolutionary processes found for this species.

Historical processes

The rapid spread of spotted knapweed on the east coast of North America, closely following the first known introduction record, was probably facilitated by the well-developed railroad network in the region (see the historical USA railroad map for 1890: Anonymous, 1951). This initial spread occurred mostly along linear transport corridors and ruderal habitats, as the colonization of natural and semi-natural habitats may have been impeded by the more competitive grassland species occurring in the eastern plains compared with the western habitats (Reinhart & Rinella, 2011). Furthermore, the spread may have been facilitated by the fact that the whole area covered by the eastern invasion is rather climatically uniform and similar to that in the introduction area and the native niche (Fig. 2; green area). Thus the species did not have to overcome significant climatic barriers to colonize the area, and hence did not require further adaptations to spread successfully.

The niche expansion characterizing the western invasion shows a remarkably different pattern. During a longer lag period, the species was confined to the Victoria coastal area (Roche *et al.*, 1986) (Figs 2a & 3c,d). The railroad network was far less developed than in the east during this period (Anonymous, 1951), making spread along transport corridors less likely. Without specific adaptation to the drier neighbouring conditions, the spread of the species outside the introduction area was probably more difficult in the west than in the east. The unintentional (probably propagules attached to undercarriages of vehicles; Sheley *et al.*, 1998) introduction to Montana around 1925 (Fig. 2) seems to have triggered its spread to most western states, which may explain the rapid expansion of all niche limits. The transition between the lag phase and the primary invasion phase occurred in the 1930–40s, with a dramatic expansion of the niche towards drier conditions, initially only in ruderal habitats. Interestingly, this rapid expansion of the niche coincided with the agricultural transition that followed the New Deal, which increased the productivity of USA agriculture by a factor of 10 and greatly modified the landscape and the environment (Conkin, 2008). These changes may have provided ample opportunities for spotted knapweed to invade open and disturbed habitats. This observed pattern suggests that human activities play an important role in driving invasion processes, by creating open niches and mediating the unintentional introduction and spread of propagules (D'Andrea *et al.*, 2009; Hufbauer *et al.*, 2012).

After about 120 years of invasion in North America, the niche limits of spotted knapweed seem to have reached a plateau for both the western and eastern invasions (Fig. 3). This suggests that the species has now colonized all types of potentially suitable climates in North America (but not necessarily all suitable places) and may thus be considered close to equilibrium with the environment. Note that the average range size of alien plants in Spain was shown to reach a maximum at 143 years (Gassó et al., 2010). Equilibrium with the environment is an important assumption when applying SDMs (Guisan & Thuiller, 2005). If true, this result suggests that previous predictions of the potential range of spotted knapweed in North America should be considered robust when based on merged data from both the native and invasive ranges (Broennimann & Guisan, 2008) and would support the use of SDM predictions to guide the management of the species (e.g. Venette *et al.*, 2010).

Ecological processes behind niche dynamics: release from biotic constraints

Grasslands are sensitive to top-down controls (e.g. through herbivory) and exhibit rapid changes in plant composition when the intensity and frequency of those controls are altered (Seastedt & Pyšek, 2011). Escaping the strongly negative competitive effects of the neighbouring vegetation with which spotted knapweed co-evolved in the native European range (Callaway *et al.*, 2011) probably favoured its establishment and spread in the invaded range. Moreover, the large-scale overgrazing of North American grasslands that started during the agricultural transition in the 1930s may have reduced local competition by native plants and further favoured the establishment of the species. Indeed, it has been shown experimentally that the presence of dominant North American native grassland species significantly reduces germination rate, seedling survivorship, growth and densities of spotted knapweed (Rinella *et al.*, 2007; Knochel *et al.*, 2010; Emery & Rudgers, 2012).

Interestingly, the negative effect of competition is greater when spotted knapweed grows with native plant species from eastern grasslands than with native species from western grasslands (Reinhart & Rinella, 2011), suggesting a biogeographically dependent response to interspecific competition. The weaker effect of native western grassland species on the growth of spotted knapweed may result from the fact that spotted knapweed is better able to adjust its photosynthesis by its ability to uptake water deeper in the soil than native grassland species (Hill *et al.*, 2006). This could contribute to the observed spatio-temporal differences of spotted knapweed invasion between eastern and western North American grasslands.

Seven seed-head and four root herbivore invertebrate species have been introduced as biological control agents to limit the spread of spotted knapweed in North America (Müller-Schärer & Schroeder, 1993). Control programmes using seed-head herbivores have been quite successful, and

the loads of these herbivore species are now comparable in both ranges, but spotted knapweed has still largely escaped the effects of root herbivores (Blair *et al.*, 2008). In addition, laboratory feeding bioassays of generalist invertebrate herbivores from the native and introduced ranges of spotted knapweed have shown that the growth of North American generalist herbivores is far lower when fed on spotted knapweed than the growth of European generalists (Schaffner *et al.*, 2011). This suggests that biogeographical differences in the response of generalist herbivores to novel plant species may have further favoured spotted knapweed invasions.

Release from biotic constraints could have been a significant driver affecting the realized niche in the new range and contributing to the observed differences between the two invasion routes. Along the western route, the niche of spotted knapweed is remarkably larger than its native niche, with realized niche limits expanded towards warmer, drier and also wetter climates. Only the cold niche limits seem to have remained stable between the native and the western part of the invaded range.

Evolutionary processes behind niche dynamics: evidence for rapid post-introduction evolution

A recent comparative demographic study has shown a pronounced increase in growth rate from European to western North American tetraploids (Hahn *et al.*, 2012), consistent with the hypothesis and experimental evidence of rapid post-introduction evolution of spotted knapweed in the invaded range (Henery *et al.*, 2010). Release from both specialist and generalist herbivores (see above) is expected to result in the evolution of increased competitive ability (the EICA hypothesis, i.e. resources devoted to defence in the native range are reallocated to growth or reproduction; Blossey & Notzold, 1995). This expectation correlates well with the findings by Broz *et al.* (2009), which showed a reduced expression of gene transcripts related to constitutive defences in introduced populations of spotted knapweed compared with native tetraploids. This evolutionary change, probably achieved during the lag phase, might then have fuelled both the fast demographic spread in the introduced range and the ability of the species to colonize natural habitats with more competitive vegetation. Preliminary chloroplast DNA analyses of the populations used in our studies have revealed that the haplotypes present in North America are similar to the most common haplotypes in Europe (U.A. Treier, Aarhus University, unpublished results), excluding the possibility of an increased growth rate in North America as a result of the spread of a haplotype that is rare in the native range. Moreover, as previous studies have provided clear evidence for multiple introductions of spotted knapweed (e.g. Marrs *et al.*, 2008), rapid evolution may be more likely than genetic drift in explaining the observed evolutionary changes (Hahn *et al.*, 2012).

Another striking result of our analyses of niche limits is the conservation of the cold niche limits across the native and invaded ranges. This pattern mirrors experimental findings for the invasive *Plantago lanceolata*, for which both native and introduced plants failed to set seed beyond the native elevational range margin and also shared a similar low-temperature limit to reproduction (Alexander *et al.*, 2012). More generally, this pattern fits the asymmetric abiotic stress limitation hypothesis (AASL), stating that species' distributions are primarily limited by physiological tolerance in the environmentally more stressful end of climatic gradients (Normand *et al.*, 2009). Stronger conservatism of the cold limit has recently been demonstrated for plant species with disjunct distributions in the Alps and the Arctic (Pellissier *et al.*, 2013). Several studies have shown that adaptations to cold climates are associated with massive reprogramming of gene expression (e.g. Survila *et al.*, 2010). Our results suggest that such adaptations have not happened in the invaded range, so that the species still occupies climatic conditions within its native fundamental (i.e. physiological) niche, but experiments are needed to substantiate this hypothesis further. Nevertheless, our study draws attention to the fact that some niche limits are more likely to be altered than others during biological invasions, depending on the prevalence of abiotic stress and competition as the main factors driving the distribution of the species (Normand *et al.*, 2009).

CONCLUSIONS

Evidence is accumulating that suggests that climatic niche shifts are occurring during biological invasions. To date, however, most studies have only assessed the overall niche change during invasions, using static approaches that compare current native and invaded ranges (e.g. Broennimann *et al.*, 2007; Fitzpatrick *et al.*, 2007; Rödder *et al.*, 2009; Medley, 2010; Petitpierre *et al.*, 2012) but do not take into account the spatio-temporal dynamics of the environmental niche along invasion routes, which is necessary to elucidate the potential mechanisms responsible for the observed changes. Our comprehensive dataset for spotted knapweed has allowed us to develop a novel approach that has revealed temporal niche dynamics along invasion routes. We have gone on to discuss these temporal patterns in the light of historical, ecological and evolutionary processes witnessed in previous studies for this species. Our results indicate that both the timing and the magnitude of ecological and evolutionary mechanisms that facilitate or prevent invasions may be intrinsically different between biogeographical regions (e.g. eastern versus western North America) and between opposite ends of ecological gradients (e.g. stressful versus competitive ends). Importantly, our study shows that detailed research on model alien invasive species, which includes both a multidisciplinary approach and a long-term commitment to data collection, offers a valuable complement to the many recent multispecies studies of niche changes during invasions. These findings significantly improve our understanding of biological invasions and may be used to direct future experimental studies to identify better the mechanisms responsible for the success of invasive species.

© 2014 John Wiley & Sons Ltd

ACKNOWLEDGEMENTS

This project was funded by the National Centre for Competence in Research (NCCR) Plant Survival, a research programme of the Swiss National Science Foundation (to A.G. and H.M.S.). We thank the curators of the herbaria who allowed us to study their collections, as well as Sarah Gray and two anonymous referees for their insightful comments on previous versions of the manuscript.

REFERENCES

Alexander, J.M., van Kleunen, M., Ghezzi, R. & Edwards, P.J. (2012) Different genetic clines in response to temperature across the native and introduced ranges of a global plant invader. *Journal of Ecology*, **100**, 771–781.

Anonymous (1951) *American railroads: their growth and development*. Association of American Railroads, Washington, DC. Maps available at: http://cprr.org/Museum/RR_Development.html (last accessed on 5 December 2013).

Blair, A.C., Schaffner, U., Hafliger, P., Meyer, S.K. & Hufbauer, R.A. (2008) How do biological control and hybridization affect enemy escape? *Biological Control*, **46**, 358–370.

Blossey, B. & Notzold, R. (1995) Evolution of increased competitive ability in invasive nonindigenous plants: a hypothesis. *Journal of Ecology*, **83**, 887–889.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, **4**, 585–589.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701–709.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481–497.

Broz, A.K., Manter, D.K., Bowman, G., Müller-Schärer, H. & Vivanco, J.M. (2009) Plant origin and ploidy influence gene expression and life cycle characteristics in an invasive weed. *BMC Plant Biology*, **9**, 33.

Callaway, R.M., Waller, L.P., Diaconu, A., Pal, R., Collins, A.R., Mueller-Schaerer, H. & Maron, J.L. (2011) Escape from competition: neighbors reduce *Centaurea stoebe* performance at home but not away. *Ecology*, **92**, 2208–2213.

Conkin, P.K. (2008) *A revolution down on the farm. The transformation of American agriculture since 1929*. University Press of Kentucky, Lexington, KY.

Crooks, J.A. (2005) Lag times and exotic species: the ecology and management of biological invasions in slow-motion. *Ecoscience*, **12**, 316–329.

Daehler, C.C. (2009) Short lag times for invasive tropical plants: evidence from experimental plantings in Hawai'i. *PLoS ONE*, **4**, e4462.

D'Andrea, L., Broennimann, O., Kozlowski, G., Guisan, A., Morin, X., Keller-Senften, J. & Felber, F. (2009) Climate change, anthropogenic disturbance and the northward range expansion of *Lactuca serriola* (Asteraceae). *Journal of Biogeography*, **36**, 1573–1587.

Di Febbraro, M., Lurz, P.W.W., Genovesi, P., Maiorano, L., Girardello, M. & Bertolino, S. (2013) The use of climatic niches in screening procedures for introduced species to evaluate risk of spread: a case with the American eastern grey squirrel. *PLoS ONE*, **8**, e66559.

Dietz, H. & Edwards, P.J. (2006) Recognition that causal processes change during plant invasion helps explain conflicts in evidence. *Ecology*, **87**, 1359–1367.

Dray, S. & Dufour, A.B. (2007) The ade4 package: implementing the duality diagram for ecologists. *Journal of Statistical Software*, **22**, 1–20.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330–342.

Emery, S.M. & Rudgers, J.A. (2012) Impact of competition and mycorrhizal fungi on growth of *Centaurea stoebe*, an invasive plant of sand dunes. *American Midland Naturalist*, **167**, 213–222.

EUNIS (2008) *European Nature Information System (EUNIS) database: habitat types and habitat classifications*. European Topic Centre on Biological Diversity, ETC/BD-EEA, Paris.

Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, **16**, 24–33.

Gallien, L., Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, **16**, 331–342.

Gassó, N., Pyšek, P., Vilà, M. & Williamson, M. (2010) Spreading to a limit: the time required for a neophyte to reach its maximum range. *Diversity and Distributions*, **16**, 310–311.

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

Gurevitch, J., Fox, G.A., Wardle, G.M. & Inderjit & Taub, D. (2011) Emergent insights from the synthesis of conceptual frameworks for biological invasions. *Ecology Letters*, **14**, 407–418.

Hahn, M., Buckley, Y. & Müller-Schärer, H. (2012) Increased population growth rate in invasive polyploid *Centaurea stoebe* in a common garden. *Ecology Letters*, **15**, 947–954.

Henery, M.L., Bowman, G., Mráz, P., Treier, U.A., Gex-Fabry, E., Schaffner, U. & Müller-Schärer, H. (2010) Evidence for a combination of pre-adapted traits and rapid adaptive change in the invasive plant *Centaurea stoebe*. *Journal of Ecology*, **98**, 800–813.

Hill, J.P., Germino, M.J., Wraith, J.M., Olson, B.E. & Swan, M.B. (2006) Advantages in water relations contribute to greater photosynthesis in *Centaurea maculosa* compared

© 2014 John Wiley & Sons Ltd

with established grasses. *International Journal of Plant Sciences*, **167**, 269–277.

Hordijk, W. & Broennimann, O. (2012) Dispersal routes reconstruction and the minimum cost arborescence problem. *Journal of Theoretical Biology*, **308**, 115–122.

Hufbauer, R.A., Facon, B., Ravigné, V., Turgeon, J., Foucaud, J., Lee, C.E., Rey, O. & Estoup, A. (2012) Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-altered habitats within the native range can promote invasions. *Evolutionary Applications*, **5**, 89–101.

Knochel, D.G., Flagg, C. & Seastedt, T.R. (2010) Effects of plant competition, seed predation, and nutrient limitation on seedling survivorship of spotted knapweed (*Centaurea stoebe*). *Biological Invasions*, **12**, 3771–3784.

Maron, J.L., Vila, M., Bommarco, R., Elmendorf, S. & Beardsley, P. (2004) Rapid evolution of an invasive plant. *Ecological Monographs*, **74**, 261–280.

Marrs, R.A., Sforza, R. & Hufbauer, R.A. (2008) Evidence for multiple introductions of *Centaurea stoebe micranthos* (spotted knapweed, Asteraceae) to North America. *Molecular Ecology*, **17**, 4197–4208.

Medley, K.A. (2010) Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus Skuse* (Culicidae), revealed by reciprocal distribution models. *Global Ecology and Biogeography*, **19**, 122–133.

Mráz, P., Bourchier, R.S., Treier, U.A., Schaffner, U. & Müller-Schärer, H. (2011) Polyploidy in phenotypic space and invasion context: a morphometric study of *Centaurea stoebe* S.L. *International Journal of Plant Sciences*, **172**, 386–402.

Mráz, P., Španiel, S., Keller, A., Bowmann, G., Farkas, A., Šingliarová, B., Rohr, R., Broennimann, O. & Müller-Schärer, H. (2012) Anthropogenic disturbance as a driver of microspatial and microhabitat segregation of cytotypes of *Centaurea stoebe* and cytotype interactions in secondary contact zones. *Annals of Botany*, **110**, 615–627.

Müller-Schärer, H. & Schroeder, D. (1993) The biological control of *Centaurea* spp. in North America: do insects solve the problem? *Pesticide Science*, **37**, 343–353.

Müller-Schärer, H., Schaffner, U. & Steinger, T. (2004) Evolution in invasive plants: implications for biological control. *Trends in Ecology and Evolution*, **19**, 417–422.

New, M., Lister, D., Hulme, M. & Makin, I. (2002) A high-resolution data set of surface climate over global land areas. *Climate Research*, **21**, 1–25.

Normand, S., Treier, U.A., Randin, C., Vittoz, P., Guisan, A. & Svenning, J.C. (2009) Importance of abiotic stress as a range-limit determinant for European plants: insights from species responses to climatic gradients. *Global Ecology and Biogeography*, **18**, 437–449.

Ochsmann, J. (2000) *Morphologische und molekularsystematische Untersuchungen an der* Centaurea stoebe *L.–Gruppe (Asteraceae–Cardueae) in Europa*. Dissertationes Botanicae, Vol. 324. J. Cramer, Berlin.

Pellissier, L., Bråthen, K.A., Vittoz, P., Yoccoz, N.G., Dubuis, A., Meier, A.S., Zimmermann, N.E., Randin, C.F., Thuiller, W., Garraud, L., Van, E.J. & Guisan, A. (2013) Thermal niches are more conserved at cold than warm limits in arctic-alpine plant species. *Global Ecology and Biogeography*, **22**, 933–941.

Peterson, A.T., Soberon, J., Pearson, R.G., Anderson, R.P., Martinez-Meyer, E., Nakamura, M. & Araújo, M.B. (2011) *Ecological niches and geographic distributions*. Princeton University Press, Princeton, NJ.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C.R., Daehler, C. & Guisan, A. (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science*, **335**, 1344–1348.

Reinhart, K.O. & Rinella, M. (2011) Comparing susceptibility of eastern and western US grasslands to competition and allelopathy from spotted knapweed *Centaurea stoebe* L. subsp. *micranthos* (Gugler) Hayek. *Plant Ecology*, **212**, 821–828.

Rinella, M.J., Pokorny, M.L. & Rekaya, R. (2007) Grassland invader responses to realistic changes in native species richness. *Ecological Applications*, **17**, 1824–1831.

Roche, B.F., Piper, G.L. & Talbott, C.J. (1986) *Knapweeds of Washington*. Washington State University, Cooperative Extension, College of Agriculture and Home Economics, Pullman, WA.

Rödder, D., Schmidtlein, S., Veith, M. & Lötters, S. (2009) Alien invasive slider turtle in unpredicted habitat: a matter of niche shift or of predictors studied? *PLoS ONE*, **4**, e7843.

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S.D., Grosberg, R.K., Hastings, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I. & Rice, W.R. (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology and Evolution*, **22**, 465–471.

Schaffner, U., Ridenour, W.M., Wolf, V.C., Bassett, T., Müller, C., Müller-Schärer, H., Sutherland, S., Lortie, C.J. & Callaway, R.M. (2011) Plant invasions, generalist herbivores, and novel defense weapons. *Ecology*, **92**, 829–835.

Seastedt, T.R. & Pyšek, P. (2011) Mechanisms of plant invasions of North American and European grasslands. *Annual Review of Ecology, Evolution, and Systematics*, **42**, 133–153.

Sheley, R.L., Jacobs, J.S. & Carpinelli, M.F. (1998) Distribution, biology, and management of diffuse knapweed (*Centaurea diffusa*) and spotted knapweed (*Centaurea maculosa*). *Weed Technology*, **12**, 353–362.

Survila, M., Heino, P. & Palva, E.T. (2010) Genes and gene regulation for low-temperature tolerance. *Genes for plant abiotic stress* (ed. by M.A. Jenks and A.J. Wood), pp. 183–219. Wiley-Blackwell, Oxford.

Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (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.

Treier, U.A., Broennimann, O., Normand, S., Guisan, A., Schaffner, U., Steinger, T. & Müeller-Schärer, H. (2009) Shift in cytotype frequency and niche space in the invasive plant *Centaurea maculosa*. *Ecology*, **90**, 1366–1377.

© 2014 John Wiley & Sons Ltd

Venette, R.C., Kriticos, D.J., Magarey, R.D., Koch, F.H., Baker, R.H.A., Worner, S.P., Raboteaux, N.N.G., McKenney, D.W., Dobesberger, E.J., Yemshanov, D., De Barro, P.J., Hutchison, W.D., Fowler, G., Kalaris, T.M. & Pedlar, J. (2010) Pest risk maps for invasive alien species: a roadmap for improvement. *BioScience*, **60**, 349–362.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Appendix S1 List of the herbaria and online databases used for collating the occurrences of spotted knapweed (*Centaurea stoebe*) in Europe and North America from 1831 to 2010.

Appendix S2 Eigenvectors (loadings) of the principal components analysis (PCA).

BIOSKETCH

Olivier Broennimann is a biogeographer interested in understanding the drivers of species' distributions under global change, in particular in the context of biological invasions. This study is part of his post-doctorate in the Guisan laboratory at the University of Lausanne.

Author contributions: O.B., A.G, P.M. and H.M.S designed the study; O.B. and P.M. collected the data; O.B. performed the analyses and drafted the manuscript; H.M.S. and P.M. strengthened the discussion on ecological and evolutionary factors; and A.G., H.M.S., P.M. and B.P. critically edited and improved the final version of the manuscript.

Editor: Richard Pearson

© 2014 John Wiley & Sons Ltd

RESEARCH PAPER

From climatic niche conservatism to spatial predictions: what can invasive bryophytes tell us?

Rubén G. Mateo[1, 2, *], Olivier Broennimann[3], Blaise Petitpierre[3], Jesús Muñoz[4], Jacques van Rooy[5], Benjamin Laenen[1], Antoine Guisan[3, 6, +], and Alain Vanderpoorten[1, +]

[1] Institute of Botany, University of Liège, 4000, Liège, Belgium, [2] Institute of Environmental Sciences, University of Castilla-La Mancha, 45071 Toledo, Spain, [3] Department of Ecology & Evolution, University of Lausanne, Biophore Building, 1015, Lausanne, Switzerland, [4] Real Jardín Botánico (CSIC), 28080, Madrid, Spain, [5] National Herbarium, South African National Biodiversity Institute, Pretoria, South Africa, [6] Institute of Earth Sciences, University of Lausanne, Geopolis Building, 1015 Lausanne, Switzerland, [+] contributed equally to this work

* Correspondence: Rubén G. Mateo, Universidad de Castilla-La Mancha, Av. Carlos III s/n, Toledo, 45071, Spain

E-mail: rubeng.mateo@gmail.com

RUNNING TITLE: Niche conservatism in alien mosses

Corresponding author: Rubén G. Mateo

Number of words in the Abstract:

Number of words in main body of the paper:

Number of references:

ABSTRACT

Aim Ecological Niche Models (ENMs) are commonly used for anticipating biological invasions but their extrapolation relies on several assumptions. Here, we test the niche conservatism hypothesis (NCH) in invasive bryophytes and investigate whether ENMs can be extrapolated to assess their invasive potential under analogous and non-analogous climates, exploring the impact of the extent of the geographical background (GB).

Location World

Methods The macroclimatic niche of *Campylopus introflexus*, *Orthodontium lineare* and *Lophocolea semiteres* was assessed with ordination techniques and partitioned into stability, unfilling, and expansion. ENMs from an ensemble model were calibrated in the native range, using ecoregions, biomes, or hemispheres as GB, and projected onto the northern hemisphere, differentiating analogous and non-analogous climate conditions.

Results No evidence for niche expansion was found. The predictive performance of the ENMs substantially differed depending on the GB but was similar in areas with analogous and non-analogous conditions. Yet, no occurrence was reported in areas with strongly non-analogous climates. The potential range of the species under analogous climate conditions was considerably larger than their extant one.

Main Conclusion The results support the NCH in bryophytes, in agreement with evidence for limited adaptive evolutionary capacities, high dispersal capacities and low levels of biotic interactions in the group. The use of large geographic units at the level of biomes or beyond is

recommended to define the extent of the GB in bryophytes and other vagile organisms with large, disjunct distributions. Although the expansion of the investigated species appears to be constrained by the presence of similar climate conditions to those currently prevailing in their native range, our analyses indicate that they might become a threat in central and eastern Europe, North America and eastern Asia if accidentally introduced.

Keywords niche modelling, niche conservatism, bryophyte, invasion, geographic background, model transferability.

INTRODUCTION

The impact of invasive species on native species, communities, and ecosystems has been widely recognized for decades (Elton, 1958; Lodge, 1993; Simberloff *et al.*, 2013). Invasive species are increasingly viewed as a significant component of global change (Vitousek *et al.*, 1996) and one of the major drivers of current biodiversity loss (Didham *et al.*, 2007). Anticipating biological invasions appears as the most effective management strategy (Leung *et al.*, 2002) and, in this regard, Ecological Niche Models (ENMs) have become the most commonly used predictive tool (Peterson, 2003; Broennimann & Guisan, 2008; Gallien *et al.*, 2010; Petitpierre *et al.*, 2012; Guisan *et al.*, 2013). ENMs are typically calibrated in the native range of species and then projected onto potential invasion ranges to identify target areas where monitoring surveys should be prioritized (Peterson, 2003; Thuiller *et al.*, 2005). The pertinence of current calibration practices of ENMs for the study of biological invasions has, however, been questioned for several reasons (Fitzpatrick *et al.*, 2007; Hill *et al.*, 2013).

First, the definition of the geographic background (GB), i.e., the extent of the study area defined to calibrate the model, has a substantial impact on the delimitation of species niches and on model transferability (Anderson & Raza, 2010; Mateo *et al.*, 2010; Barve *et al.*, 2011; Acevedo *et al.*, 2012). Models calibrated in small GBs are likely to underestimate the importance of large-scale factors such as climate (Jiménez-Valverde *et al.*, 2011) and to include a lower number of variables (Van der Wal *et al.*, 2009). Models built in large GBs may, in turn, unduly down-weight the importance of fine-scale conditions (Lobo *et al.*, 2010).

Second, forecasting potential distributions of species outside of their native range may be problematic because the observed distribution of a species alone provides no information about how the species might respond under novel environments (Fitzpatrick & Hargrove, 2009). Failure to identify regions with non-analogous environments can result in

misinterpretation of potential ranges in invasive species, as ENMs may identify regions as highly suitable simply due to inappropriate extrapolation of response curves (Fitzpatrick & Hargrove, 2009).

Third, ENMs rely on the hypothesis of niche conservatism (NCH) in space and time (Pearman *et al.*, 2008; Gallien *et al.*, 2012; Petitpierre *et al*., 2012). The introduction of different lineages that do not necessarily occur in sympatry in their native range may, however, allow lineages to cross, potentially resulting in increased fitness and ecological tolerance (Ellstrand & Schierenbeck, 2000; Mukherjee *et al.*, 2012). Additionally, release from biotic interactions can make alien species more competitive in the invaded range (the evolution of increased competitive ability hypothesis, Blossey & Nötzold, 1995). The NCH has therefore been recurrently challenged (e.g., Broennimann *et al*., 2007; Gallagher *et al*., 2010; Medley, 2010; Hill *et al*., 2012; Stigall, 2012; Erfmeier, 2013; Hill *et al*., 2013; Hulme & Barrett, 2013; but see Alexander *et al*., 2012; Petitpierre *et al*., 2012; Strubbe *et al.*, 2013), and is further complicated by its scale dependence (Mandle *et al.*, 2010; Petersen, 2013).

Application of the NCH is expected to hold true in groups of organisms with either limited adaptive evolutionary capacities, high dispersal capacities so that gene flow cancels adaptations in niche traits, and/or low levels of biotic interactions (Cooper *et al*., 2010). These conditions are typically met in bryophytes, a group of spore-producing plants wherein both experimental (Lönnell *et al*., 2012, 2013; Sundberg, 2013) and genetic evidence (Piñeiro *et al.*, 2012; Szövényi *et al.,* 2012) point to the high long-distance dispersal capacities. The poor digestibility of bryophyte tissues and the potential presence of deterrents (see Vanderpoorten & Goffinet, 2009, for review, but see Maciel-Silva & Dos-Santos, 2011; Fang & Zhu, 2012) result in remarkably low direct consumption. This, along with the low levels of competitive exclusion among bryophyte communities (Steel *et al*., 2004, but see Ingerpuu & Vellak,

2013), suggests that the extent to which bryophyte communities are shaped by biotic interactions is limited.

In bryophytes, genetic variation does not appear to be adaptive (see Shaw & Goffinet, 2000, for review, but see Szövényi *et al*., 2009; Hutsemékers *et al*., 2010; Pisa *et al*., 2013). In the expanding moss *Pogonatum dentatum* for example, shoots transplanted from other sites or areas reached a taller size than native ones, indicating that local adaptation has not occurred (Hassel *et al*., 2005b). By contrast with flowering plants, selection of broadly adapted, 'all-purpose' genotypes in bryophytes suggests that physiological acclimatization and pre-adaptation to a range of environmental conditions is much more important than is genetic specialization (Shaw & Goffinet, 2000).

In Europe, three among the 22 introduced species, namely the southern hemisphere *Campylopus introflexus, Orthodontium lineare*, and *Lophocolea semiteres*, have spread in the course of the 20th century (Söderström, 1992; Stieperaere, 1994; Hassel & Söderström, 2005). Although the number of invasive bryophytes species strongly pales by comparison with seed plants, they threaten habitats that are often species rich and of high conservation relevance (Klinck, 2010). In some instances, a thick bryophyte cover may directly prevent the regeneration of the herb and shrub layers. In relatively dry environments, such as grasslands and heathlands, even a thin moss cover may inhibit seedling emergence by increasing the time above-ground of dispersed seeds, thus increasing the chances of destruction by fire and predation, and by creating a physical barrier to germination for seeds (Zamfir, 2000). Dense carpets of *C. introflexus* may therefore reduce the germination of the heath, *Calluna vulgaris*, up to 60%, potentially threatening its ability to regenerate (Equihua & Usher, 1993). In coastal dunes, *C. introflexus* can form dense carpets that largely replace native vegetation and

affect habitat structure and ground-dwelling arthropod species and functional diversity (Schirmel *et al.*, 2011; Schirmel & Buchholz, 2013).

By comparing the macroclimatic niches of *C. introflexus*, *O. lineare* and *L. semiteres* in their areas of origin (southern hemisphere; hereafter native range) and in the range they have invaded (northern hemisphere; hereafter invaded range), the primary aim of the present study was to test the following hypotheses:

1. If bryophytes indeed exhibit a low potential for local adaptation and if their communities are hardly shaped by biotic interactions, we expect that they conserve their niche through space and time. Within the climatic envelop of analogous climates in the native and invaded ranges, we therefore do not expect species to occupy portions of their niche in the invaded range that would not be also occupied in the native one (niche expansion sensu Petitpierre *et al.*, 2012).
2. If bryophyte species are pre-adapted to a broad range of climatic conditions but that large parts of these conditions do not exist and thus are not occupied by the species in their southern native ranges due to the scarcity and different configuration of landmasses, we expect that they may occur in the invaded range under climatic conditions that may not be analogous to those observed in the native range.

We then assess whether ENMs developed in the southern hemisphere can be projected to the northern hemisphere to assess their invasive potential, exploring the impact of the extent of the GB on models and their transferability.

MATERIAL AND METHODS

Model species

Campylopus introflexus, *O. lineare* and *L. semiteres* were first observed in Europe in 1941, 1910, and 1965, respectively (Stieperaere, 1994; Hassel & Söderström, 2005). *Campylopus introflexus* has further been reported from western North America since 1975 (Klinck, 2010) and was most recently reported as adventive in eastern North America (Miller, 2009). Both *C. introflexus* and *O. lineare* are equipped with spores of less than 20µm (Hassel & Söderström, 2005, Klinck, 2010), a critical condition for wind long-distance dispersal (Wilkinson *et al.*, 2012). *Campylopus introflexus* further disperses by vegetative diaspores that are assumed to contribute to local spread (Klinck, 2010). In the dioicous *L. semiteres*, the two sexes exhibit distinct distribution patterns so that sporophyte production is extremely rare in the invasion range (Stieperaere, 1994), potentially limiting its dispersal capacity.

Campylopus introflexus exhibits a wide ecological tolerance (Klinck, 2010). *Orthodontium lineare* occurs on decaying wood under light tree cover (Hedenäs *et al*., 1989). *Lophocolea semiteres*, which was primarily thought to be restricted to acidic sand areas (Stieperaere, 1994), was subsequently found under different soil conditions in disturbed areas (Vanderpoorten, 1997). While the importance of such ecological factors cannot be dismissed to explain the distribution patterns of those species at the local scale, they do not, however, appear to determine their range at the continental scale.

Distribution data and environmental predictors

The total number of occurrences recorded for *C. introflexus*, *L. semiteres* and *O. lineare* across their entire native and invasion range was 42,064, 2,222 and 13,646, respectively. These occurrences were recorded from databased and georeferenced herbarium specimens and further completed with a review of the literature (Appendix S1 in Supporting

Information). To avoid sampling bias (Syfert *et al.*, 2013), only points that were separated by at least 0.16 decimal degrees from each other (i.e., matching the resolution of the climatic data) were retained, resulting in 1,271 data points (483 in the native range, and 788 in the invaded range) for *C. introflexus*; 351 (258 in the native range, and 93 in the invaded range) for *L. semiteres;* and 579 (123 in native range, and 456 in invaded range) for *O. lineare* (Fig. 1).

Thirty-five variables from CliMond (Kriticos *et al.*, 2012) as well as monthly and annual potential evapotranspiration, (downloaded from the CGIAR-CSI website, http://www.cgiar-csi.org, Zomer *et al.*, 2008) were employed as environmental predictors. Following Barve *et al.* (2011) and Acebedo *et al.* (2012), biogeographic units were employed to circumscribe the extent of the geographic background (GB) where these variables were measured. Three spatial extents, namely ecoregions and biomes, as defined by Olson *et al*. (2001), and hemispheres, were used in this study. We mapped the ecoregions, biomes, and hemispheres, where each species was reported (Fig. 1), and sampled 10,000 pixels at 10 minutes resolution as background data (or pseudo-absences). To avoid multicollinearity, we ran a correlation analysis between each pair of variables and eliminated the ones with a Pearson correlation value>0.8. The final set of variables used to run the models were: temperature seasonality, mean temperature of warmest quarter, moisture index seasonality, mean moisture index of warmest quarter, and annual potential evapotranspiration.

Statistical analyses

Ordinations

A Principal Component Analysis (PCA) of the climate conditions recorded at each background point was performed to identify the combinations of variables that best capture macroclimatic variation. We calibrated PCA's at each background level (ecoregions, biomes and hemispheres respectively) to assess the impact of the background on the measure of niche shifts. The first two PCA axes jointly accounted for about 80% of the total macroclimatic variance for each species when ecoregions, biomes, and hemispheres were employed as a GB, respectively. PCA1 was negatively (positively in *L. semiteres*) correlated with the mean temperature of warmest quarter (BIO10), moisture index seasonality (BIO31), and potential evapotranspiration in *C. introflexus* and *O. lineare*, and positively correlated (negatively in *L. semiteres*) with temperature seasonality (BIO04) and mean moisture index of warmest quarter (BIO34). PCA2 was for all species positively correlated with temperature seasonality and negatively with the mean moisture index of warmest quarter. Following Broennimann *et al.* (2012), the actual observations were then plotted and transformed into densities in a gridded environmental space depicted by the first two PCA axes. Then, we partitioned the niche of each species into three components according to Petitpierre *et al* (2012), namely stability S (the proportion of the niche that is occupied in the invaded range and shared with the native range); unfilling U (the proportion of the niche in the native range that is not occupied in the invaded range); and expansion E (the proportion of the niche in the invaded range that is not occupied in the native range). To avoid that niche differences would reflect climatic differences between the native and invaded backgrounds, these statistics were computed on the climate space shared between the two ranges (i.e. in analogous climates sensu Veloz *et al*., 2012).

Ecological niche models (ENMs)

Species niches were characterized using an ensemble model (Araújo & New, 2007) of four different techniques: generalized linear models (GLM; McCullagh & Nelder, 1989), MaxEnt (Phillips *et al.*, 2006), gradient boosting machine (GBM; Friedman, 2001) and Random forests (RF; Breiman, 2001). We used the BIOMOD 2.0 package in R (Thuiller *et al.*, 2009, www.R-Forge.R-project.org) for the modelling with GBM, BRT and GLM with all the default parameters. For MAXENT, the regularization multiplayer was changed to the value of two to avoid overprediction. For each technique, presences and pseudo-absences used to calibrate the model were weighted such as to ensure neutral (0.5) prevalence.

ENMs tend to perform better in the native range when both native and invaded range data are available for model calibration (Broennimann & Guisan, 2008; Gallien *et al.*, 2010; Medley, 2010). Since the focus of the present study was not to produce the best-fit model in the native range, but rather to determine the performance of the models to predict invasions, we calibrated the models with all the data from the native range and employed the presences in the invaded range to evaluate the capacity of ENMs to predict invasions.

For each species, consensus models of the four employed techniques were generated (see Appendix S2 and S3 for an assessment of the performance of each technique and of the consensus models in the native range, respectively). The performance of the consensus models in the invaded range, i.e., their capacity to predict invasions, was measured with the Boyce index (Hirzel *et al.*, 2006) based on observed presences in the invaded range (Petitpierre *et al.*, 2012).

Maps of the potential presence of each species in their invasion range were finally generated. For that purpose, the continuous suitability index was transformed into a binary presence/absence variable, using a threshold maximizing both sensitivity and specificity in the native range (see Appendix S4 for model sensitivity and specificity).

Since model transferability was questioned under non-analogous conditions (Fitzpatrick & Hargrove, 2009), the potential invasion range and model performance (as measured by the Boyce index) were assessed under both analogous and analogous+non-analogous conditions. Analogous conditions were defined from an analysis of the climatic similarity between the native and invasion ranges using multivariate environmental similarity surfaces (MESS; Elith *et al*., 2010). MESS measures the similarity of any given pixel in the invaded range to a reference set of pixels in the native range with respect to the chosen predictor variables. A pixel with a positive value indicates that it falls within the range of environmental values present in the native range, while a pixel with a negative value indicates that at least one variable has a value that is outside of the range of environmental values present in the native range. Since negative values close to zero could mean that the climatic conditions are only slightly different from the native range, we defined analogous climatic conditions with three different threshold: 0, -10 and -20.

The entire procedure was repeated for each of the three GBs, using the 10,000 pseudo-absences generated within each of the ecoregions, biomes, and hemisphere where the species occurs in its native range. The impact of the different background strategies on the extent of the predicted invasion range was assessed by calculating the mean correlations (Pearson coefficient) between all consensus models and by measuring the mean overlap between all consensus binary models depending on the GB (see Appendix S6).

RESULTS

Niche comparisons

The stability of the macroclimatic niche of *C. introflexus, L. semiteres* and *O. lineare* in their invaded range was very high, with a minimum of S=0.93. None of these species largely

expanded their niche during the invasion process, E ranging between 0.00 and 0.07 depending on the extent of the GB considered (Fig. 2, Appendix S5). The proportion of the niche occupied in the native range but not in the invaded range (unfilling) increased with the level of spatial extent of the GB and ranged between 0.14 in *O. lineare* and 0.67 in *L. semiteres* at the level of ecoregions, 0.35 in *O. lineare* and 0.68 in *L. semiteres* at the level of biomes, and between 0.73 in *O. lineare* and 0.91 in *L. semiteres* at the level of the biosphere.

Ecological niche models

The potential areas of invasion of *C. introflexus*, *O. lineare* and *L. semiteres* as inferred from the consensus binary models for each GB and projected under both analogous and non-analogous climate conditions in the invaded range are presented in Fig. 3. Globally, very similar values of the Boyce index were observed when considering analogous only or analogous+non-analogous conditions (Table 1 and Appendix S3). 29, 19 and 2 % of the observed occurrences in the invasion range fell outside of the area characterized by analogous climate conditions (MESS=0) for *C. introflexus*, *O. lineare* and *L. semiteres*, respectively (Appendix S7). All of the actual observations were reported, however, from areas with sub-analogous climate conditions (MESS threshold of -10 and -20, see yellow and green areas in Fig. 3b), and no occurrence was observed in areas with strongly non-analogous climates (MESS threshold <-20, red area in Fig. 3b).

Substantial differences in the performance of the consensus models when predicting the occurrence in the invasion range, as measured by the Boyce index, were observed depending on the spatial extent of the GB (Table 1). With *L. semiteres* and *O. lineare*, a substantially higher Boyce index was obtained from models employing biomes to define the extent of the GB. In *C. introflexus*, the highest Boyce index was observed with models employing either

ecoregions or the southern hemisphere to define the GB. These differences in model performance in terms of ability to predict invasion were paralleled by substantial differences in the extent of the potential area of invasion (Fig. 3, see also Appendix S6 for the mean percentage overlap between binary models and mean correlation Pearson coefficient for original consensus models).

With the best-fit models either employing ecoregions or the southern hemisphere to define the extent of the GB in *C. introflexus*, the potential range of the species under non-analogous climate conditions is much larger in Europe and in North America than its extant one, especially in the eastern central region where the species is currently absent. The analyses also identify a large potentially suitable area in eastern Asia, where the species does not currently occur. When the models are projected in areas with analogous climates only, the same trend is observed, but the spatial extent of the potential areas identified in North America, Europe, and Asia, is much reduced, so that 29% of the actual occurrences in Europe are not included (Fig 3b). Similar trends can be observed in *L. semiteres* and *O. lineare*, with a large suitability across Europe that largely exceeds the current distribution and large potential areas in Eastern Central America, Central and Eastern Asia where the species have not yet been reported.

DISCUSION

No clear evidence for macroclimatic niche expansion was found in the invaded range of the three investigated bryophyte species, whatever the extent of the geographic background. This is the first piece of evidence pointing to niche conservatism in bryophytes, in agreement with our first initial hypothesis based on evidence for limited adaptive evolutionary capacities,

high dispersal capacities and low levels of biotic interactions. The present results further parallel recent studies supporting the idea that widespread alien species do globally not expand their niche in the invasion area (Alexander *et al.*, 2012; Petitpierre *et al.*, 2012; Strubbe *et al*., 2013). This finding has important consequences for the application of ENMs and their transferability in the increasingly growing number of studies employing ENMs in bryophytes (Kruijer *et al*., 2010; Hutsemékers *et al*., 2011; Sérgio *et al*., 2011; Delgadillo *et al*., 2012; Mateo *et al.*, 2013; Patiño *et al.*, 2013; Yu *et al*., 2013), and especially for their use as 'canaries in the coal mine' in the current context of climate change (Tuba *et al*., 2011; Désamoré *et al.*, 2012).

While the proportion of niche expansion in the invaded range was not affected by the extent of the GB, the proportion of the niche occupied in the native range but not in the invaded one (unfilling) substantially increased with increasing GBs. Similarly, the predicted size of the invaded range under analogous+non-analogous climate conditions was inversely proportional to the size of the biogeographic units employed to define the GB in the native range. The same trends were observed by Acevedo *et al.* (2012), who further suggested that models built from large GBs exhibit high discriminatory power but are hardly informative. Anderson & Raza (2010) suggested that large GBs make the models prone to overfit the environmental conditions present in the region occupied by the species. This may happen because the algorithm recognizes spurious environmental differences between the inhabited localities and localities that could be inhabited but are not due dispersal limitations or historical events restricting the species current distribution. Yet, models using larger biogeographic regions such as biomes, or in some instances whole hemispheres, to define the

GB, performed better than models using smaller geographic units (regions) in their ability to predict invasions, as shown by their substantially higher Boyce index.

The three species considered in this study in fact display a highly disjunct distribution across the southern hemisphere, with a range scattered across southern South America, South Africa, and Australia. This is a typical trend in bryophyte species that tend to display highly disjunct and larger distributions than seed plant species (Vanderpoorten *et al.*, 2010). For instance, 43% of the species of mosses that are found in North America are also found in Europe while 70% of the moss species found in Europe also occur in North America (Frahm & Vitt, 1993). By contrast, 48% of genera, but only 6.5% of the species of vascular plants, are shared between the North American and European floras (Qian, 1999). As the GB should not only reflect the extant but also potentially occupied range in the past (Acevedo *et al.*, 2012), the use of large geographic units at the level of biomes or beyond is therefore recommended to define the extent of the GB in bryophytes and other vagile organisms with large, disjunct distributions.

The projection of the models in the northern hemisphere under analogous and analogous+non-analogous climate conditions resulted in a very similar Boyce index. The projection of the models under analogous+non-analogous climate conditions in the northern hemisphere resulted in a potential range that is considerably larger than the realized one for the three investigate species. This range encompasses large portions of central and eastern Europe, North America and eastern Asia, where these species are still rare or absent. Although *L. semiteres* largely fails to reproduce sexually in Europe, the two other species are well equipped for efficient long-distance dispersal (see above). In the expanding moss *Pogonatum dentatum*, higher genetic structuring in the area of origin than in the area of invasion points to enhanced spore-mediated gene flow in the latter (Hassel *et al*., 2005a),

paralleling similar evidence for increased dispersal capacities of alien species in the invasion range (Ward *et al*., 2012, but see Colautti *et al*., 2012). In the peatmoss *Sphagnum subnitens*, the spread across North America by a single genotype points to very fast colonization rates (Karlin *et al.,* 2011). Dispersal limitations are therefore unlikely to explain the discrepancy between the potential and realized range in invasive bryophytes. Rather, while the distribution of the three investigated species in the invaded range was not restricted to areas with strictly analogous climates (MESS threshold >0), all of the actual observations in the invasion range were reported from areas with sub-analogous climates (MESS threshold >-20). This suggests that, in agreement with our second initial hypothesis, those species are, to some extent, pre-adapted to a range of climatic conditions, but are unable to invade areas characterized by substantially different climates than those currently prevailing in their native range. Comparative distribution maps through time indicate that these species are still expanding their range in Europe (Hassel & Söderström, 2005). Since the macroclimatic niche of the three investigated species is compatible with conditions that prevail in eastern Asia and North America under analogous or nearly analogous conditions, they might follow a similar invasion pattern as in Europe if accidentally introduced, calling for a targeted monitoring program aiming at erasing any source population at the very beginning of the invasion process.

ACKNOWLEDGEMENTS

RGM and AV gratefully acknowledge financial support from the Belgian Funds for Scientific Research (FNRS) (grants 1.5036.11 and 2.4557.11) and the University of Liège (Grant C 11/32). AV also acknowledges support from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement ES-TAF-1699 (SYNTHESYS).

REFERENCES

Acevedo, P., Jiménez-Valverde, A., Lobo, J.M. & Real, R. (2012) Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*, **39**, 1383-1390.

Alexander, J.M., van Kleunen, M., Ghezzi, R. & Edwards, P.J. (2012) Different genetic clines in response to temperature across the native and introduced ranges of a global plant invader. *Journal of Ecology*, **100**, 771-781.

Anderson, R.P. & Raza, A. (2010) The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *Journal of Biogeography*, **37**, 1378-1393.

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42-47.

Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810-1819.

Blossey, B. & Nötzold, R. (1995) Evolution of increased competitive ability in invasive nonindigenous plants: a hypothesis. *Journal of Ecology*, **83**, 887-889.

Breiman, L. (2001) "Random Forests". *Machine Learning*, **45**, 5-32.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, **4**, 585-589.

Broennimann, O., Treier, U.A., Müller-Schärer, H., Thuiller, W., Peterson, A.T. & Guisan, A. (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701-709.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481-497.

Colautti, R.I., White, N.A. & Barrett, S.C.H. (2012) Variation of self-incompatibility within invasive populations of purple loosestrife (*Lythrum salicaria* L.) from eastern north America. *International Journal of Plant Sciences*, **171**, 158–166.

Cooper, N., Jetz, W. & Freckleton, R.P. (2010) Phylogenetic comparative approaches for studying niche conservatism. *Journal of Evolutionary Biology*, **23**, 2529–2539.

Delgadillo, C., Villaseñor, J.L. & Ortiz, E. (2012) The potential distribution of *Grimmia* (Grimmiaceae) in Mexico. *The Bryologist*, **115**, 12–22.

Désamoré, A., Laenen, B., Stech, M., Papp, B., Hedenäs, L., Mateo, R.G. & Vanderpoorten, A. (2012) How do temperate bryophytes face the challenge of a changing environment? Lessons from the past and predictions for the future. *Global Change Biology*, **18**, 2915-2924.

Didham, R.K., Tylianakis, J.M., Gemmell, N.J., Rand, T.A. & Ewers, R.M. (2007) Interactive effects of habitat modification and species invasion on native species decline. *Trends in Ecology and Evolution*, **22**, 489-496.

Elith, J., Kearney, M. & Phillips, S. (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330-342.

Ellstrand, N.C. & Schierenbeck, K.A. (2000) Hybridization as a stimulus for the evolution of invasiveness in plants? *Proceedings of the National Academy of Sciences*, **97**, 7043-7050.

Elton, C.S. (1958) *The Ecology of Invasions by Animals and Plants*. Chicago University.

Equihua, M. & Usher, M.B. (1993) Impact of carpets of the invasive moss *Campylopus introflexus* on *Calluna vulgaris* regeneration. *Journal of Ecology*, **81**, 359-365.

Erfmeier, A. (2013) Constraints and release at different scales – The role of adaptation in biological invasions. *Basic and Applied Ecology*, **14**, 281-288.

Fang, Y. & Zhu, R.-L. (2012) *Haplocladium microphyllum* (Hedw.) Broth. capsules as food for *Agrotis* sp. (Lepidoptera) larvae. *Journal of Bryology*, **34**, 108-113.

Fitzpatrick, M.C. & Hargrove, W.W. (2009). The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation*, **18**, 2255-2261.

Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict its native range? *Global Ecology and Biogeography*, **16**, 24-33.

Frahm, J.-P. & Vitt, D.H. (1993) A comparison of the mossfloras of Europe and North America. *Nova Hedwigia*, **56**, 307–333.

Friedman, J.H. (2001) Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, **29**, 1189-1232.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology*, **98**, 790-799.

Gallien, L., Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, **16**, 331-342.

Gallien, L., Douzet, R., Pratte, S., Zimmermann, N.E. & Thuiller, W. (2012) Invasive species distribution models – how violating the equilibrium assumption can create new insights. *Global Ecology and Biogeography*, **21**, 1126-1136.

Hassel, K. & Söderström, L. (2005) The expansion of the alien mosses *Orthodontium lineare* and *Campylopus introflexus* in Britain and continental Europe. *Journal of the Hattori Botanical Laboratory*, **97**, 183-193.

Hassel, K., Såstad, S. M., Gunnarsson, U. & Söderström, L. (2005a) Genetic variation and structure in the expanding bryophyte *Pogonatum dentatum* in its area of origin and a newly colonised area. *American Journal of Botany*, **92**, 1684–1690.

Hassel, K., Pedersen, B. & Söderström, L. (2005b) Changes in life history traits in an expanding species, phenotypic plasticity and genetic differentiation: a reciprocal transplantation experiment with *Pogonatum dentatum*. *Ecography,* **28**, 71–80.

Hedenäs, L., Herben, T., Rydin, H. & Söderström, L. (1989) Ecology of the invading moss species *Orthodontium lineare* in Sweden: Spatial distribution and population structure. *Ecography*, **12**, 163-172.

Hill, M.P., Chown, S.L. & Hoffmann, A.A. (2013) A predicted niche shift corresponds with increased thermal resistance in an invasive mite, *Halotydeus destructor*. *Global Ecology and Biogeography*, **22**, 942-951.

Hill, M.P., Hoffmann, A.A., Macfadyen, S., Umina, P.A. & Elith, J. (2012) Understanding niche shifts: using current and historical data to model the invasive redlegged earth mite, *Halotydeus destructor*. *Diversity and Distributions*, **18**, 191-203.

Hirzel, A.H., Le Lay, G., Helfer, V., Randin, C. & Guisan, A. (2006) Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, **199**, 142-152.

Hulme, P.E. & Barrett, S.C.H. (2013) Integrating trait- and niche-based approaches to assess contemporary evolution in alien plant species. *Journal of Ecology*, **101**, 68-77.

Hutsemékers, V., Hardy, O.J., Mardulyn, P., Shaw, A.J. & Vanderpoorten, A. (2010) Macroecological patterns of genetic structure and diversity in the aquatic moss *Platyhypnidium riparioides*. *New Phytologist*, **185**, 852-864.

Hutsemékers, V., Szoveńyi, P., Shaw, A.J., Gonzaĺez-Mancebo, J.M., Munõz, J. & Vanderpoorten, A. (2011) Oceanic islands are not sinks of biodiversity in spore-producing plants. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 18989-18994.

Ingerpuu, N. & Vellak, K. (2013) Growth depends on neighbours: experiments with three *Sphagnum* L. species. *Journal of Bryology*, **35**, 27-32.

Jiménez-Valverde, A., Lira-Noriega, A., Barve, N., Maher, S.P., Nakazawa, Y., Papes, M., Soberon, J., Sukumaran, J. & Peterson, A.T. (2011) Dominant climate influences on North American birds. *Global Ecology and Biogeography*, **20**, 114–118.

Karlin, E.F., Andrus, R.E., Boles, S.B. & Shaw, A.J. (2011) One haploid parent contributes 100% of the gene pool for a widespread species in northwest North America *Molecular Ecology*, **20**, 753–767.

Klinck, J. (2010): NOBANIS – Invasive Alien Species Fact Sheet – *Campylopus introflexus*. – From: Online Database of the European Network on Invasive Alien Species - NOBANIS www.nobanis.org, Date of access 1/9/2013.

Kriticos, D.J., Webber, B.L., Leriche, A., Ota, N., Macadam, I., Bathols, J. & Scott, J.K. (2012) CliMond: global high resolution historical and future scenario climate surfaces for bioclimatic modelling. *Methods in Ecology and Evolution*, **3**, 53-64.

Kruijer, J.D., Raes, N. & Stech, M. (2010) Modeling the distribution of the moss species *Hypopterygium tamarisci* (Hypopterygiaceae, Bryophyta) in Central and South America. *Nova Hedwigia*, **91**, 399–420.

Leung, B., Lodge, D.M., Finnoff, D., Shogren, J.F., Lewis, M.A. & Lamberti, G. (2002) An ounce of prevention or a pound of cure: bioeconomic risk analysis of invasive species. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **269**, 2407-2413.

Lobo, J.M., Jime´nez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103–114.

Lodge, D.M. (1993) Biological invasions: Lessons for ecology. *Trends in Ecology and Evolution*, **8**, 133-137.

Lönnell, N., Hylander, K., Jonsson, B.G. & Sundberg, S. (2012) The fate of the missing spores — patterns of realized dispersal beyond the closest vicinity of a sporulating moss. *PLoS ONE*, **7**, e41987.

Lönnell, N., Jonsson, B.G. & Hylander, K.. (2013) Production of diaspores at the landscape level regulates local colonization: an experiment with a spore-dispersed moss; *Ecography*, in press.

Maciel-Silva, A.S. & Dos-Santos, N.D. (2011) Detecting herbivory in two mosses from an Atlantic Forest, Brazil *Journal of Bryology*, **33**, 140-147.

Mandle, L., Warren, D.L., Hoffmann, M.H., Peterson, A.T., Schmitt, J. & von Wettberg, E.J. (2010) Conclusions about niche expansion in introduced *Impatiens walleriana* populations depend on method of analysis. *PLoS ONE*, **5**, e15297.

Mateo, R.G., Croat, T.B., Felicísimo, Á.M. & Muñoz, J. (2010) Profile or group discriminative techniques? Generating reliable species distribution models using pseudo-absences and target-group absences from natural history collections. *Diversity and Distributions*, **16**, 84-94.

Mateo, R.G., Vanderpoorten, A., Muñoz, J., Laenen, B. & Désamoré, A. (2013) Modeling species distributions from heterogeneous data for the biogeographic regionalization of the European bryophyte flora. *PLoS ONE*, **8**, e55648.

McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*, 2nd edn. Chapman & Hall, London.

Medley, K.A. (2010) Niche shifts during the global invasion of the Asian tiger mosquito, *Aedes albopictus* Skuse (Culicidae), revealed by reciprocal distribution models. *Global Ecology and Biogeography*, **19**, 122-133.

Miller, N.G. (2009) Mosses adventive and naturalized in the northeastern United States: New examples and new distributional records. *Rhodora*, **111**, 218-230.

Mukherjee, A., Williams, D.A., Wheeler, G.S., Cuda, J.P., Pal, S. & Overholt, W.A. (2012) Brazilian peppertree (*Schinus terebinthifolius*) in Florida and South America: evidence of a possible niche shift driven by hybridization. *Biological Invasions*, **14**, 1415-1430.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P. &

Kassem, K.R. (2001) Terrestrial ecoregions of the world: a new map of life on Earth. *Bioscience*, **51**, 933-938.

Patiño, J., Medina, R., Vanderpoorten, A., González-Mancebo, J.M., Werner, O., Devos, N., Mateo, R.G., Lara, F. & Ros, R.M. (2013) Origin and fate of the single-island endemic moss *Orthotrichum handiense*. *Journal of Biogeography*, **40**, 857-868.

Petersen, M. (2013) Evidence of a climatic niche shift following North American introductions of two crane flies (Diptera; genus *Tipula*). *Biological Invasions*, **15**, 885-897.

Peterson, A.T. (2003) Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology*, **78**, 419-433.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science*, **335**, 1344-1348.

Phillips, S.J., Anderson, R.P. & Schapire, R.P. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.

Piñeiro, R., Popp, M., Hassel, K., Listl, D., Westergaard, K.B., Flatberg, K.I., Stenøien, H.K., Brochmann, C. (2012) Circumarctic dispersal and long-distance colonization of South America: The moss genus *Cinclidium*. *Journal of Biogeography*, **39**, 2041-2051.

Pisa, S., Werner, O., Vanderpoorten, A., Magdy, M. & Rosa, R.M. (2013) Altitudinal genetic differentiation of the cosmopolitan moss *Bryum argenteum* in the mountains of Sierra Nevada (Spain, SW Europe). *American Journal of Botany*, **100**, 2000-2008.

Qian, H. (1999) Spatial pattern of vascular plant diversity in North America north of Mexico and its floristic relationship with Eurasia. *Annals of Botany*, **83**, 271-283.

Schirmel, J. & Buchholz, S. (2013) Invasive moss alters patterns in life-history traits and functional diversity of spiders and carabids. *Biological Invasions*, **15**, 1089-1100.

Schirmel, J., Timler, L. & Buchholz, S. (2011) Impact of the invasive moss *Campylopus introflexus* on carabid beetles (Coleoptera: Carabidae) and spiders (Araneae) in acidic coastal dunes at the southern Baltic Sea. *Biological Invasions*, **13**, 605–620.

Sérgio, C., Figueira, R. & Menezes, R. (2011) Modeling the distribution of *Sematophyllum substrumulosum* (Hampe) E. Britton as a signal of climatic changes in Europe. *Bryophyte Ecology and Climate Change* (ed. by Z. Tuba, N.G. Slack and L.R. Stark), pp. 427-440. Cambridge University Press, Cambridge.

Shaw, A.J. & Goffinet, B. (2000) *Bryophyte Biology*. Cambridge University Press.

Simberloff, D., Martin, J.-L., Genovesi, P., Maris, V., Wardle, D.A., Aronson, J., Courchamp, F., Galil, B., GarcÌa-Berthou, E., Pascal, M., Pyaek, P., Sousa, R., Tabacchi, E. & Vila, M. (2013) Impacts of biological invasions: what's what and the way forward. *Trends in Ecology and Evolution,* **28**, 58-66.

Söderström, L. (1992). Invasions and range expansions and contractions of bryophytes. *Bryophytes and Lichens in a Changing Environment* (ed. by J.W. Bates and A.M. Farmer), pp. 131-158. Clarendon Press, Oxford.

Steel, J.B., Wilson, J.B., Anderson, B.J., Lodge, R.H.E. & Tangney, R.S. (2004) Are bryophyte communities different from higher-plant communities? Abundance relations. *Oikos*, **104**, 479-486.

Stieperaere, H. (1994) *Lophocolea semiteres* (Lehm.) Mitt. in Belgium and The Netherlands, another antipodal bryophyte spreading on the European continent. *Lindbergia*, **19**, 29-36.

Strubbe, Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography*, in press.

Stigall, A.L. (2012) Using ecological niche modelling to evaluate niche stability in deep time. *Journal of Biogeography*, **39**, 772-781.

Sundberg, S. (2013) Spore rain in relation to regional sources and beyond. *Ecography*, **36**, 364-373.

Syfert, M.M., Smith, M.J. & Coomes, D.A. (2013) The effects of sampling bias and model complexity on the predictive performance of MaxEnt species distribution models. *PLoS ONE*, **8**, e55158.

Szövényi, P., Hock, Zs., Korpelainen, H. & Shaw, A.J. (2009) Spatial pattern of nucleotide polymorphism indicates molecular adaptation in the bryophyte *Sphagnum fimbriatum*. *Molecular Phylogenetics and Evolution*, **53**, 277-286.

Szövényi, P., Sundberg, S. & Shaw, A.J. (2012) Long-distance dispersal and genetic structure of natural populations: an assessment of the inverse isolation hypothesis in peat mosses. *Molecular Ecology*, **21**, 5461-5472.

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

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

Tuba, Z., Slack, N.G. & Stark, L.R. (2011) *Bryophyte Ecology and Climate Change*. Cambridge University Press, Cambridge.

Vanderpoorten A. (1997) A bryological survey of the Brussels Capital Region. *Scripta Botanica Belgica*, **14**, 1-83.

Vanderpoorten, A. & Goffinet, B. (2009) *Introduction to Bryophytes*. Cambridge University Press, Cambridge.

Vanderpoorten, A., Gradstein, S.R., Carine, M.A. & Devos, N. (2010) The ghosts of Gondwana and Laurasia in modern liverwort distributions. *Biological Reviews*, **85**, 471-487.

Van der Wal, J., Shoo, L.P., Graham, C. & Williams, S.E. (2009) Selecting pseudo-absence data for presence-only distribu- tion modeling: how far should you stray from what you know? *Ecological Modelling*, **220**, 589–594.

Veloz, S.D., Williams, J.W., Blois, J.L., He, F., Otto-Bliesner, B. & Liu, Z. (2012) No-analog climates and shifting realized niches during the late quaternary: implications for 21st-century predictions by species distribution models. *Global Change Biology*, **18**, 1698–1713.

Vitousek, P.M., D'Antonio, C.M., Loope, L.L. & Westbrooks, R. (1996) Biological invasions as global environmental change. *American Scientist*, **84**, 468.

Ward, M., Johnson, S. & Zalucki, M. (2012) Modes of reproduction in three invasive milkweeds are consistent with Baker's Rule. *Biological Invasions*, **14**, 1237-1250.

Wilkinson, D. M., Koumoutsaris, S., Mitchell, E. A. D. & Bey, I. (2012) Modelling the effect of size on the aerial dispersal of microorganisms. *Journal of Biogeography*, **39**, 89–97.

Yu, J., Ma, Y.H. & Guo, S.L. (2013) Modeling the geographic distribution of the epiphytic moss *Macromitrium japonicum* in China. *Annales Botanici Fennici*, **50**, 35–42.

Zamfir, M. (2000) Effects of bryophytes and lichens on seedling emergence of alvar plants: evidence from greenhouse experiments. *Oikos*, **88**, 603–611.

Zomer, R.J., Trabucco, A., Bossio, D.A. & Verchot, L.V. (2008) Climate change mitigation: A spatial analysis of global land suitability for clean development mechanism afforestation and reforestation. *Agriculture, Ecosystems & Environment*, **126**, 67-80.

BIOSKETCH

Rubén G. Mateo focuses on biogeography, spatial ecology, plant ecology, spatial biodiversity patterns, conservation biology, effects of alien species and climate change on biodiversity. He is working on the optimization of different methods and techniques to produce reliable species distribution models. He is also involved in floristic studies in the Iberian Peninsula.

TABLES

Table 1. Boyce index of the consensus models from four ENMs calibrated on the 100% distribution of *Campylopus introflexus*, *Orthodontium lineare*, and *Lophocolea semiteres* in their native range (southern hemisphere) and projected onto their invasion range (Europe) under both analogous (bold) and non-analogous climate conditions (in parentheses) depending on the level of spatial extension of the geographic background (ecoregions, biomes and hemispheres).

SPECIES	*C. introflexus*	*L. semiteres*	*O. lineare*
BACKGROUND			
Ecoregions	**0.968** (0.935)	**0.881** (0.721)	**-0.583** (0.029)
Biomes	**0.699** (0.777)	**0.909** (0.914)	**0.902** (0.951)
Hemispheres	**0.954** (0.975)	**0.732** (0.735)	**0.802** (0.835)

FIGURE LEGENDS

Figure 1. Occurrences used to model the macroclimatic niche of *Campylopus introflexus*, *Lophocolea semiteres*, and *Orthodontium lineare* in their area of origin and invasion (left). The distribution of the species at the level of ecoregions and biomes is depicted in the central and right columns, respectively.

Figure 2. Partitioning of the macroclimatic niche of *Campylopus introflexus*, *Lophocolea semiteres*, and *Orthodontium lineare* in their area of origin (southern hemisphere) and area of invasion (Europe) into (i) stability (proportion of the niche in the invasive range shared with the native range, in blue), (ii) unfilling (proportion of the niche occupied in the area of origin but not in the area of invasion, in green), and (iii) expansion (proportion of the niche occupied in the area of invasion but not in the area of origin, or 1-stability, in red) in the first plane of a Principal Component Analysis of the macroclimatic conditions found at the levels of ecoregions, biomes, and hemispheres, where the species occurs. The solid and dashed contour lines illustrate, respectively, 100% and 50% of the available (background) environment in the areas of origin (green) and invasion (red). The arrows represent how the centre of the niche has changed between the native and invasive areas.

Figure 3. Potential area of invasion of the southern hemisphere bryophyte species *C. introflexus*, *O. lineare* and *L. semiteres* as inferred from the consensus model of four ENM techniques calibrated in the southern hemisphere (native range) and projected in the northern hemisphere (invasion range) depending on the spatial extent of the geographic background. The models were projected under both analogous and analogous+non-analogous climate

conditions as compared to the native range, as defined by a MESS analysis using different thresholds (0, -10 and -20). (a) World scale. (b) European scale for selected GB strategies. Blue dots represent actual datapoints.

.

FIGURES

Figure 1

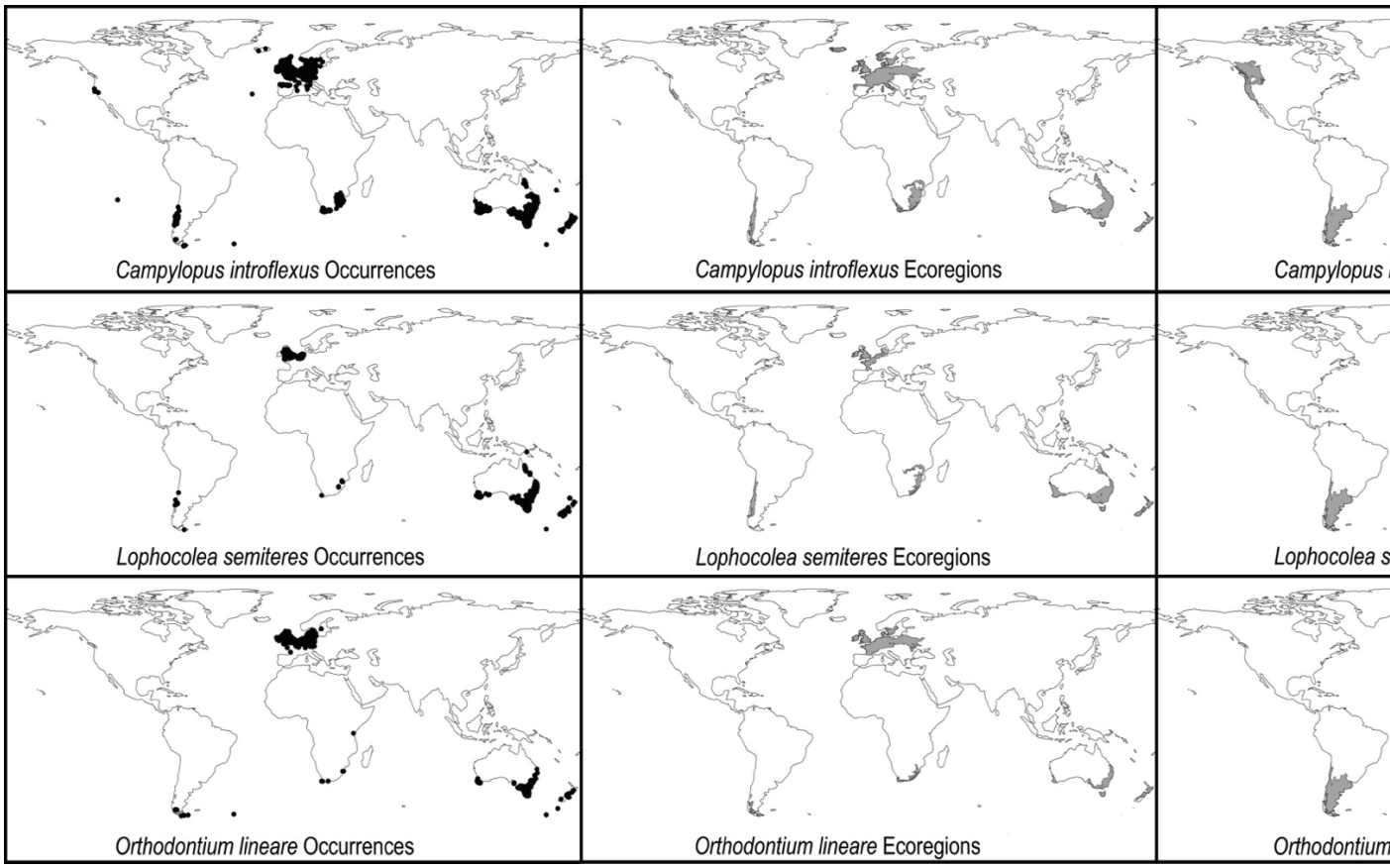


Figure 2

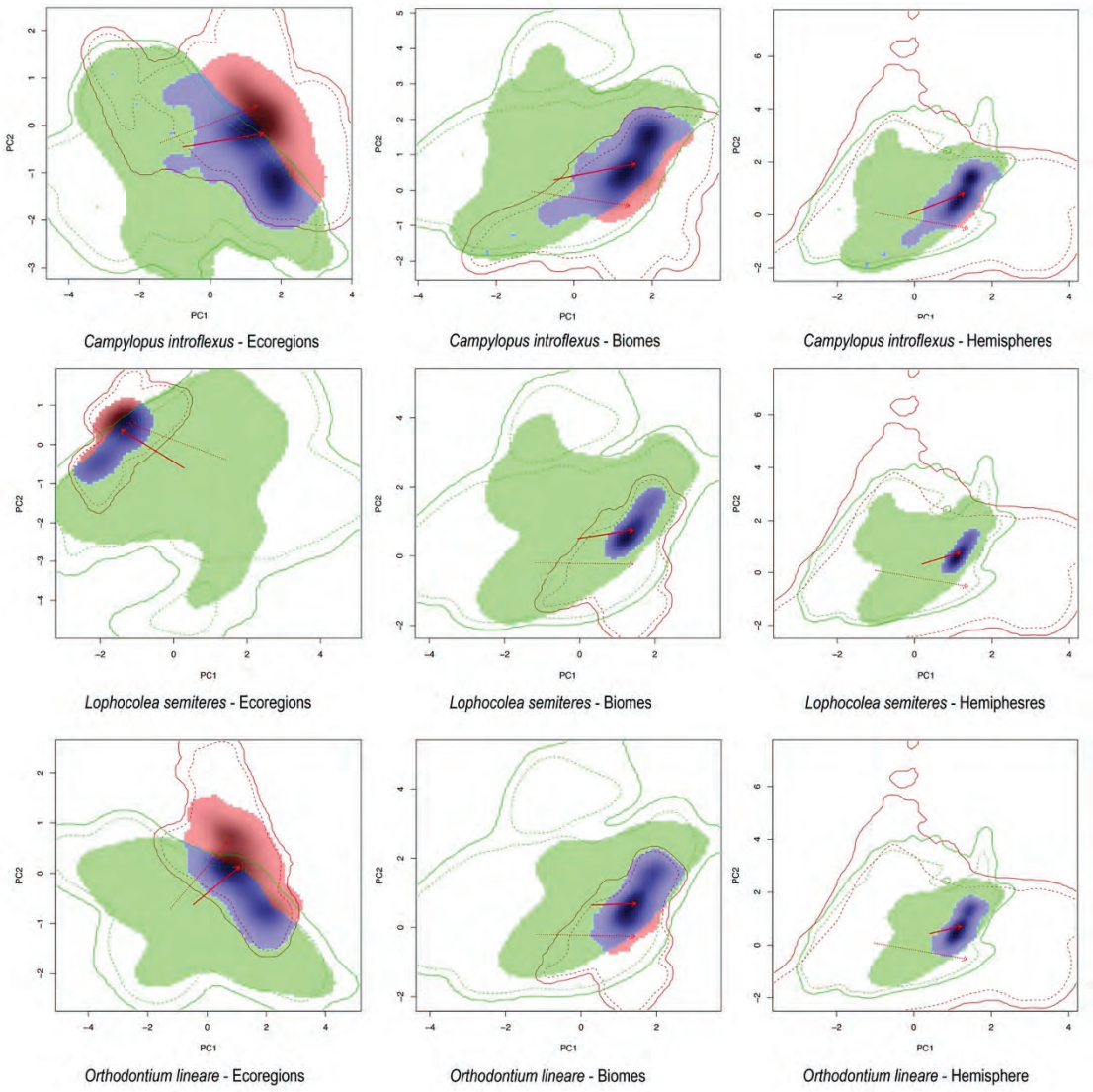

PC1
PC2
Campylopus introflexus - Ecoregions
Campylopus introflexus - Biomes
Campylopus introflexus - Hemispheres
Lophocolea semiteres - Ecoregions
Lophocolea semiteres - Biomes
Lophocolea semiteres - Hemiphesres
Orthodontium lineare - Ecoregions
Orthodontium lineare - Biomes
Orthodontium lineare - Hemisphere

Figure 3a

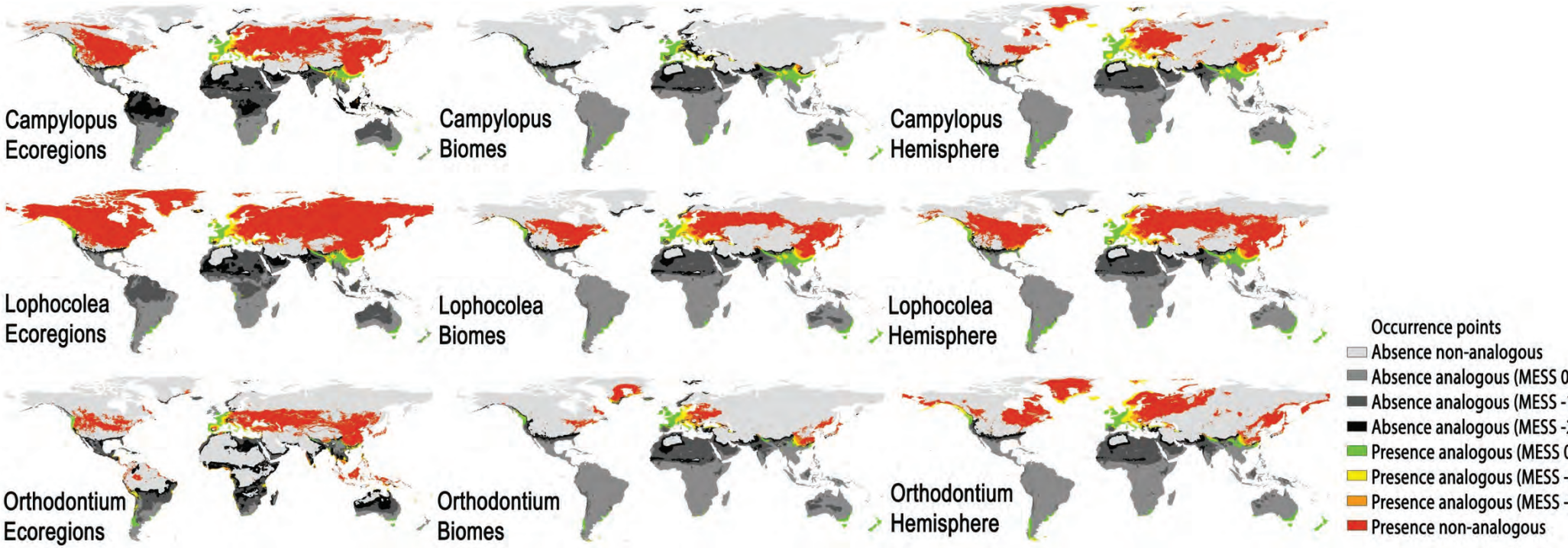
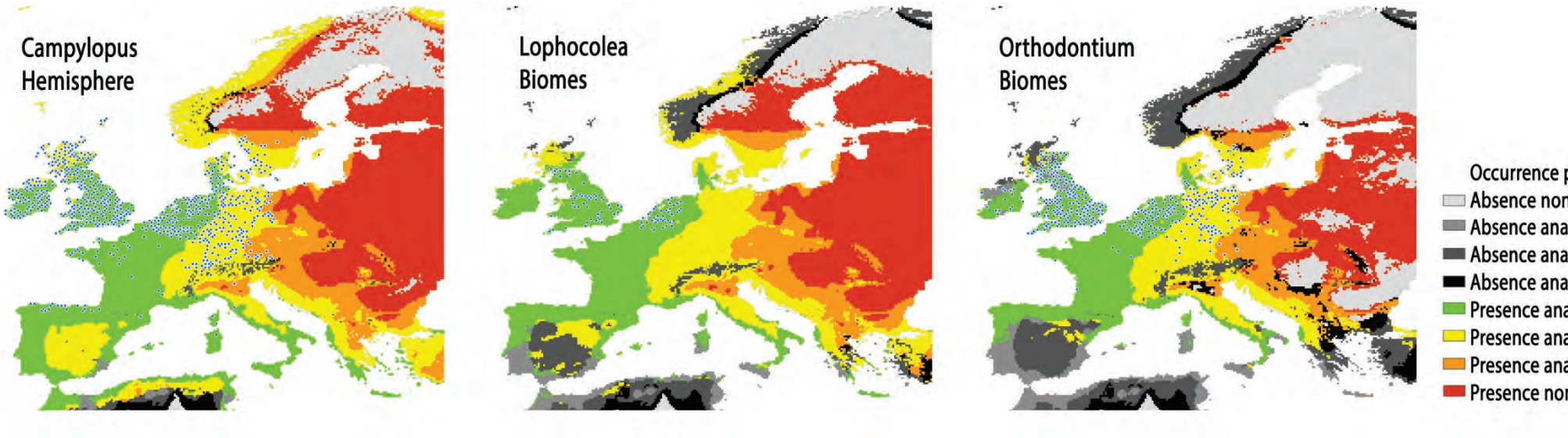

Campylopus
Ecoregions
Campylopus
Biomes
Campylopus
Hemisphere
Lophocolea
Ecoregions
Lophocolea
Biomes
Lophocolea
Hemisphere
Orthodontium
Ecoregions
Orthodontium
Biomes
Orthodontium
Hemisphere
Occurrence points
Absence non-analogous
Absence analogous (MESS 0
Absence analogous (MESS -
Absence analogous (MESS -
Presence analogous (MESS 0
Presence analogous (MESS -
Presence analogous (MESS -
Presence non-analogous

Figure 3b


Campylopus
Hemisphere
Lophocolea
Biomes
Orthodontium
Biomes
Occurrence p
Absence non
Absence ana
Absence ana
Absence ana
Presence ana
Presence ana
Presence ana
Presence non

SUPPORTING INFORMATION

Appendix S1. Source information of the distribution data employed to build Ecological Niche Models in the native range of three invasive bryophyte species (*Campylopus introflexus* and *Orthodontium lineare* and the liverwort *Lophocolea semiteres*) in the northern hemisphere.

Appendix S2. Performance of the four ENM techniques employed to model the macroclimatic niche of *C. introflexus*, *O. lineare* and *L. semiteres* in their native range.

Appendix S3. Predictive performance of the consensus models of four ENMs used to model the macroclimatic niche of *C. introflexus*, *O. lineare* and *L. semiteres* in their native and invaded range.

Appendix S4. Contrasting the performance of AUC and TSS as criteria to optimize the binary models.

Appendix S5. Partitioning of the macroclimatic niche of *Campylopus introflexus*, *Lophocolea semiteres*, and *Orthodontium lineare* in their area of origin (southern hemisphere) and are of invasion (Europe) into (i) stability (i.e., proportion of the niche in the invasive range shared with the native range), (ii) unfilling (i.e., proportion of the niche occupied in the area of origin but not in the area of invasion), and (iii) expansion (i.e., proportion of the niche occupied in the area of invasion but not in the area of origin, or 1-stability) depending on the level of spatial extension (ecoregions, biomes and hemispheres) used to sample the background points.

Appendix S6. Mean Pearson correlation coefficient and mean percentage overlap (+/- SD) between the potential range of the southern hemisphere invasive bryophyte species *C. introflexus, O. lineare* and *L. semiteres* as inferred from original and binary consensus models respectively of four ENM techniques depending on the extent of the geographic background employed to calibrate the models in the native range.

Appendix S7. Number of observed occurrences and percentage in the invasion range that felt in analogous climates conditions and outside of the area characterized by analogous climate conditions. The non-analogous climate area was generated with three different threshold (0, -10 and -20).

Measuring the relative effect of factors affecting species distribution model predictions

Emeric Thibaud,[1*] Blaise Petitpierre,[2] Olivier Broennimann,[2]
Anthony C. Davison,[1] and Antoine Guisan[2,3]

[1]Chair of Statistics, Ecole Polytechnique Fédérale de Lausanne,
EPFL-FSB-MATHAA-STAT, Station 8, 1015 Lausanne, Switzerland
[2]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland
[3]Faculty of Geoscience and the Environment, University of Lausanne, 1015 Lausanne, Switzerland

[*]To whom correspondence should be addressed; E-mail: emeric.thibaud@epfl.ch.

Summary

1. Species distribution models are increasingly used to address conservation questions, so their predictive capacity requires careful evaluation. Previous studies have shown how individual factors used in model construction can affect prediction. Although some factors probably have negligible effects compared to others, their relative effect sizes are largely unknown.

2. We introduce a general virtual ecologist framework to study the relative importance of factors involved in the construction of species distribution models.

3. We illustrate the framework by examining the relative importance of five key factors—namely, a missing covariate, spatial autocorrelation in presence-absence occurrences from a dispersal process, sample size, sampling design and modeling technique—in a real study framework based on plants in a mountain landscape at regional scale, and show that, for the parameter values considered here, most of the variation in prediction accuracy is due to sample size and modeling technique. Contrary to repeatedly reported concerns, the presence of spatial autocorrelation has only comparatively small effects.

4. This study shows the importance of using a nested statistical framework to evaluate the relative effects of various factors that may affect species distribution models.

Key-words: linear mixed-effects model, relative importance, spatial autocorrelation, species distribution model, virtual ecologist

Introduction

Prediction based on species distribution models (SDMs) has grown rapidly (Guisan & Thuiller, 2005; Elith & Leathwick, 2009; Franklin, 2010; Peterson *et al.*, 2011), stimulated by the pressing need identified a decade ago for more predictive models (Clark *et al.*, 2001; Côté & Reynolds, 2002) and by their use in global change assessments (e.g. Schröter *et al.*, 2005; Pereira *et al.*, 2010). SDMs are empirical models based on observed occurrences or abundances of species. By fitting the observed environmental conditions in which a species can grow and survive, these models are rooted in the realized environmental niche concept, i.e., the envelope of environmental conditions constrained by competitive interactions and dispersion limitations (Guisan & Thuiller, 2005; Soberón, 2007).

Many methodological decisions must be made during the construction of an SDM (Guisan & Thuiller, 2005; Franklin, 2010), introducing various possible sources of uncertainty for model building and predictions (Beale & Lennon, 2012), and it is important to understand these. Some factors can be partly controlled (e.g., sample size, sampling design, modeling technique), but others are harder to handle (e.g., species dispersal and species distribution patterns). Most previous studies have assessed the effects of factors on SDM building and predictions only one at a time (Table S5, but see Dormann *et al.*, 2008; Diniz-Filho *et al.*, 2009; Garcia *et al.*, 2012), thus preventing the calculation of the relative effects for the different factors. Most such studies have used real data, but a few have used simulated, or so-called virtual, data (Zurell *et al.*, 2010; Miller, 2014).

Studies using real data have mostly proceeded by degrading initial conditions individually—e.g., adding spatial autocorrelation, coarsening the grain or decreasing sample size—and measuring the resulting increase in prediction error and decrease in predictive power. Even if much has been learned from them, these single-factor analyses cannot allow us to gauge the relative effects of different factors on SDM outcomes. Other studies have assessed how variation in these factors may affect future predictions (Thuiller *et al.*, 2004; Baselga & Araújo, 2009; Buisson *et al.*, 2010), but are uninformative about the relative effect sizes because the truth is unknown and only uncertainty can be quantified. Calculating the relative effect sizes of several factors in model building requires a different approach in which the truth is known. Using a virtual ecologist approach (Zurell *et al.*, 2010; Miller, 2014) with simulated virtual species distribution data and virtual observers offers such a truth-proofing perspective, and allows full control of how factors are combined. No consensus exist yet in the literature on how to best simulate virtual species distribution. Proposed approaches include the fit of an initial model (e.g., a generalized linear model, GLM, McCullagh & Nelder, 1989) to real data (as in Wisz & Guisan, 2009; Albert *et al.*, 2010), defining theoretical response functions (e.g., Hirzel & Guisan, 2002) or using more dynamic approaches (e.g., Zurell *et al.*, 2010) and then considering the resulting spatial predictions as the true distributions.

The goal of this paper is to suggest an approach to systematically assess how the factors involved in the construction of an SDM affect its prediction performance, using virtual species simulations. Our approach is based on environmental covariates that reflect natural situations, and is illustrated by a comparison of the effects of sample size, sampling design, modeling technique, and spatial autocorrelation resulting from a dispersal process or a missing covariate. Unlike previous studies, the present simulations include all possible combinations of the factors, and use replication to assess the variability due to different presence-absence and sampling patterns. Our approach can easily be extended to other factors and situations.

Materials and methods

We decompose our simulation scheme into three parts, corresponding to three critical steps of species distribution modeling. The first step is the simulation of virtual species, with or without a missing covariate, and with or without spatial autocorrelation in the simulated presences and absences. In practice the choice of species is dictated by the study goals, and is not under the control of the modeler. The second step is to sample these virtual species at different locations, controlling the sample size and the sampling design. In practice this step is determined by the resources available for data gathering. The third step is to apply different modeling techniques to the data obtained at the sampling step, and to evaluate their performances using the root mean squared error. The technique is chosen by the modeler, but in practice its performance cannot be directly evaluated, since no ground truth is available. Our approach is implemented in version 3.0.1 of the open-source software R (R Core Team, 2013). The code is given in the Supporting Information.

Virtual species simulation

Our simulation uses a real Alpine landscape in the western Swiss Alps ($7°2'$–$7°14'$ E; $46°28'$–$46°31'$ N). This landscape represents an area of about 700 km^2 of the canton of Vaud, comprising $1,127,599$ pixels at a resolution size of 25 meters. It has been intensively sampled (Pottier *et al.*, 2013), and this helps in providing realistic estimates of parameters needed for the simulation. We used five real climatic and topographic predictors, labelled $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_5$: degree days above three (*ddeg300*), a moisture index between June and August (*mmind68*), daily average global potential shortwave radiation per month (*sumradyy*), the annual average number of frost days during the growing season (*sfroyy*) from existing maps (Zimmermann & Kienast, 1999), and topographic position (*topos*, Randin *et al.*, 2006). Figure 1 shows the spatial patterns of these predictors.

We generated virtual species distributions (see Figure 1) based on $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_5$. As no consensus exist yet in the literature on how to best simulate virtual species distribution, we used an approach based on fitting a GLM to real data, projecting it throughout the study area and then considering the resulting spatial predictions as the true distributions (see Wisz & Guisan, 2009; Albert *et al.*, 2010). Let $\boldsymbol{X}$ denote the $N \times 5$ matrix of predictors, each row of which corresponds to a different pixel. To create a surface $\{p_1, \ldots, p_N\}$ of presence probabilities from the matrix $\boldsymbol{X}$, we used the probit function to relate the predictors to the presence probabilities, through the expression

$$p(\boldsymbol{X}) = \Phi\{\eta(\boldsymbol{X})\}, \tag{1}$$

where Φ is the standard Gaussian cumulative distribution function. For simplicity we included only quadratic terms for each predictor and we did not consider interactions, i.e., we let

$$\eta(\boldsymbol{X}) = \alpha_0 + \sum_{j=1}^{5} \alpha_j (\boldsymbol{x}_j - \beta_j)^2, \tag{2}$$

with coefficients α_j and β_j chosen to produce a realistic distribution for each virtual species. We used different values of α and β to construct $S = 10$ surfaces, thereby creating S virtual species with different responses to the predictors. These species were generated using parameters estimated by fitting model (1) to data on ten real species distributed in the study area, chosen to represent

different taxonomic groups, distributions and habitat types (see Supporting Information). Once a presence probability function $p(\boldsymbol{X})$ corresponding to a virtual species was defined, we simulated a presence or absence by generating independent standard normal variables $\gamma_1, \ldots, \gamma_N$, as then the random variables

$$Y_i = \begin{cases} 1, & \gamma_i \leq \eta_i, \\ 0, & \gamma_i > \eta_i, \end{cases} \tag{3}$$

are equal to 1 with probabilities $p_i = \Phi(\eta_i)$ given by Equation (1).

We used the probit link in (1) because its representation (3) in terms of normal variables simplified the simulation of correlated presences-absences. The difference with a logit model was shown to be small enough not to affect our results.

Spatial autocorrelation (SAC) represents the clustering tendency of presences and absences (Cablk *et al.*, 2002; Segurado *et al.*, 2006; Dormann, 2007; Dormann *et al.*, 2008). It may be explained by the effect of spatial covariates such as altitude or by a biological process of species dispersal or interactions with other species. Although these are two different types of SAC, they both usually result in spatially correlated residuals from the fit of a SDM, and it is difficult to distinguish them based on data (Guisan & Thuiller, 2005; Dormann *et al.*, 2007; Beale *et al.*, 2010). The distinction between these types of SAC is important for the simulation: SAC due to missing covariates is present at the probability level and is not supposed to vary with replications, whereas SAC due to a biological process appears only at the presence-absence level and has no effect on marginal probabilities p_i. We simulated these two types of SAC as follows.

SAC due to an unobserved spatially varying covariate was generated by excluding a predictor from our virtual dataset, so that the spatial variation of the presences could not be entirely explained using the remaining predictors. For each virtual species, the predictor to be excluded was determined by fitting a probit model (1) to the simulated presence-absence data, measuring the reduction in the likelihood due to excluding one predictor at a time from Equation (2), and then excluding the second most important predictor. Correlation between the spatial predictors means that this type of SAC will generally alter the estimates of the remaining parameters. The range of SAC for the excluded variable is not controlled.

We also simulated SAC representing a stochastic colonization or dispersal process, where species presences tend to cluster, by using correlated γ_i in (3): we replaced the independent γ_i by a spatial Gaussian process $\gamma(\boldsymbol{s})$, where $\boldsymbol{s}$ is the pixel location. The process $\gamma(\boldsymbol{s})$ has zero mean, unit variance and correlation function ρ; see Figure 1. This type of SAC is easy to implement and its range can be precisely controlled using the parameters of the correlation function. The marginal probability of presence at location $\boldsymbol{s}$ is still defined from Equation (1), but with presences and absences tending to cluster. Thus this type of SAC increases the variance of SDM predictions but does not bias them. We used the correlation function $\rho(h) = \exp(-h/\lambda)$, where h is the distance between locations in kilometers and λ is a range parameter. We estimated λ by fitting a spatial probit model to the ten real species that were used to create our virtual species, giving estimates of λ that were all smaller than 0.5 km (see Supporting Information, Table S1). Although we cannot be sure that this SAC corresponds to a dispersal process, we fixed $\lambda = 0.5$ in our simulation for comparability with the real data. This is justified by the plausible short-range dependence in plant species distributions; here the correlation of $\gamma(\boldsymbol{s})$ is lower than 0.05 beyond 1.5 kilometers, an upper bound for the maximum distance reported for most dispersal strategies among the plant species in the study area described by Vittoz & Engler (2007).

For each of the $S = 10$ virtual species, and for each combination of missing covariate (true or

false) and dispersal (true or false), we independently simulated the spatial presence-absence process $R = 10$ times to account for stochastic variability.

Sampling virtual species

When sampling virtual species, we controlled the sample size n and the design. We created training samples with $n = 100$ and $n = 500$, sampling both presences and absences. The mean proportion of presences (i.e., the prevalence) in the training samples for each species varied from 31% to 59%. We used two sampling designs: a simple random (or uniform) design (Albert *et al.*, 2010), for which every pixel has the same probability of being sampled; and a design under which locations that are close to roads are more likely to be sampled, as occurs frequently in practice (also called sampling bias, see Graham *et al.*, 2004); see Figure 1. With each training sample of size n we also uniformly sampled an independent test sample $\mathcal{T}$ of size $n' = 5000$ from all presences and absences, except those of the training sample, which is reserved for the evaluation of the fitted models.

This sampling of the virtual species is repeated $R' = 5$ times on each of the $R = 10$ previously simulated presences and absences for each of the $S = 10$ simulated species. This hierarchical simulation scheme permits us to separate the effects of species variation and sampling procedure. For computational reasons, simulation of the dispersal process takes place during the random sampling stage, as it is infeasible to simulate the Gaussian process at 10^6 sites.

Fits and evaluations

We used four different modeling techniques to estimate the presence-absence probabilities for each simulated species, using five or four predictors depending on whether some SAC was due to a missing predictor. The first was a generalized linear model (GLM, McCullagh & Nelder, 1989) for binary responses (the presence-absence data) with the probit link function, quadratic terms for each predictor and no interactions. If no covariate is missing, this is the true model that generates our data via Equation (1). The second, which was fitted using the R package gam (Hastie, 2013), was a generalized additive model (GAM, Hastie & Tibshirani, 1990) with binary response, the probit link, smoothing spline terms and no interactions between predictors. The third was a maximum entropy (MaxEnt, Phillips *et al.*, 2006) model fitted using the R package dismo (Hijmans *et al.*, 2013) with default settings. MaxEnt is a technique that models presence-only data using background data (Elith *et al.*, 2011; Renner & Warton, 2013; Merow *et al.*, 2013). Renner & Warton (2013) show that MaxEnt is equivalent to Poisson regression, and link it to a Poisson point process model, thus giving insight into how MaxEnt deals with presence-only data. Because absence data were available to us, we used MaxEnt in a non-standard manner, using the absence data in place of a background sample. We used the logistic output of MaxEnt (Yackulic *et al.*, 2013). Finally, random forests (RF, Breiman, 2001) were fitted (using the R package randomForest, Liaw & Wiener, 2002) with default settings, i.e., 500 trees and with two variables at each node of each tree. Each of these four modelling techniques provides estimates $\hat{p}_i$ at the locations $i \in \mathcal{T}$, which, under some general SDM assumptions (Guisan & Thuiller, 2005; Elith & Leathwick, 2009; Araújo & Peterson, 2012) and algorithm-specific assumptions (Hastie & Fithian, 2013), estimate the true presence probabilities p_i.

There are many measures of model accuracy (Fielding & Bell, 1997; Caruana & Niculescu-Mizil, 2004; Liu *et al.*, 2011). We focused on the root mean squared error (RMSE, Caruana & Niculescu-Mizil, 2004; Liu *et al.*, 2011) owing to its advantages in our context: it compares probabilities to probabilities, it is easily interpretable, and the distribution of log(RMSE) was found to be appropriate for the analysis of variance. Other measures, such as the area under the receiver operating characteristic curve (AUC, Mason & Graham, 2002) or the point-biserial correlation (COR, Tate, 1954), popular in SDM, can also be computed, but since[1] their distributions are difficult to model and interpret, further analysis is more complicated. The RMSE is computed by comparison of the estimated presence probabilities with the true marginal probabilities (unknown in real applications):

$$\text{RMSE} = \sqrt{\frac{1}{n'} \sum_{i \in \mathcal{T}} (\hat{p}_i - p_i)^2},$$

with $n' = \text{Card}(\mathcal{T})$, and where the true probabilities p_i come from Equation (1). This RMSE corresponds to the accuracy related to the estimation of the environmental niche of the species, excluding possible undesirable effects of the SAC at the presence-absence level. The RMSE may be interpreted as the mean distance between predicted and true probabilities over the locations of the test sample. Use of independent test samples, i.e., pixels that are not in the training samples, avoids bias in estimation of prediction accuracy, and because it compares predictions with true presence probabilities, and not observed presences or absences, it is unaffected by any proximity of the pixels used for the training and test samples. As the RMSE involves the true p_i, it can be computed in simulation settings, though not for real data. More discussion may be found in the Supporting Information.

SDMs can typically be evaluated on other landscapes than those used to fit the models (Randin *et al.*, 2006), where different interactions between the predictors can affect their prediction accuracy. The procedure presented here easily extends to validation in external landscapes. In addition to measuring the importance of the factors in the original landscape, we projected the true distributions of the ten virtual species into three other regions of Switzerland, sampled test localities and calculated the RMSE in these regions (see Supporting Information).

Statistical analysis

The simulation scheme described above requires the application of four modeling techniques to each of $2^4 \times R \times R' \times S$ samples. With $S = 10$ simulated species, $R = 10$ presence-absence simulations, and $R' = 5$ sampling patterns, we have 8000 datasets, to each of which we apply four techniques. The entire procedure, from the simulation of species to the evaluation of predictions, takes less than a day using the statistical environment R with parallel computing on eight cpus. Although there is no limitation on the number of presence-absence simulations R or the number of species S, it is more difficult to increase the number R' of sampling patterns, as the simulation of dispersal using the Gaussian process involves the calculation of inverse matrices of side approximately $1200 \times R'$; this can be computer-intensive.

In order to assess the importance of the different factors we propose to study variation in the values of log(RMSE) via a linear mixed-effects model. The log transformation stabilises the variance of the RMSEs; see Figure 2. For each simulated species there are five factors: *missing* for the

[1] Antoine suggere d'ajouter ici: "they require binarizing the initial probabilistic distribution and because"

presence of a missing influential covariate (T/F); *dispersal* for the presence of a dispersal process (T/F); *n*, the sample size (100 or 500); *design* for the sampling design (simple random/road-based); and finally *technique* for the different modeling techniques.

Since we built our simulation hierarchically, values from the same species, from the same random simulation or from the same sampling pattern are expected to be similar. Hence we use a model with three nested random effects, corresponding to the different species, the random simulation and the random sampling pattern:

$$\log(\text{RMSE})_{ksinbjt} = \alpha_0 + \alpha_{m*s*n*d*t} + \varepsilon_k + \varepsilon_{kmsi} + \varepsilon_{kmsindj} + \varepsilon_{kmsindjt}, \tag{4}$$

with α_0 the intercept and $\alpha_{m*s*n*d*t}$ the 32 fixed-effect parameters for *missing* (m), *dispersal* (s), *n* (n), *design* (d), *technique* (t) and all the interactions of all orders. The indices $k = 1, \ldots, 10$, $i = 1, \ldots, 10$ and $j = 1, \ldots, 5$ respectively represent the ten simulated species, the ten presence-absence simulations and the five sampling patterns. The ε_k, ε_{kmsi} and $\varepsilon_{kmsindj}$ are independent zero-mean random variables that account for the species random effects, the presence-absence simulations and the different sampling patterns. The $\varepsilon_{kmsindjt}$ are independent centered normal random variables whose variance σ^2 corresponds to the finest level of variation, representing the errors in the linear mixed-effects model. This is a classical split-plot design with nested random effects; see Section 10.2 of Venables & Ripley (2002).

Different procedures have been proposed to examine the relative importances of factors in linear models. Hierarchical partitioning (Chevan & Sutherland, 1991; Mac Nally, 2000) has been widely used, and in a similar context by Dormann *et al.* (2008) in particular. In a balanced design such as ours, this method essentially compares the sums of squares corresponding to each factor in the analysis of variance (ANOVA, Table 1). Although the importances of different factors are easily compared when the sums of squares for interactions are tiny compared with those for main effects, this is harder when some interactions have large effects, as in the present case. Thus, in addition to ANOVA, we also use the marginal and conditional coefficients of determination, R^2, that were defined in the context of mixed-effects models by Nakagawa & Schielzeth (2013). We use the R package nlme (Pinheiro *et al.*, 2013) to fit model (4) by restricted maximum likelihood (REML, Venables & Ripley, 2002), with each factor and all its interactions excluded, and compare the resulting marginal R^2 with that of the full model. Thus, excluding one factor leads to the exclusion of 16 terms, allowing us to measure the full contribution of that factor to the model (4), including any interactions with other factors.

Results

The RMSE for the original landscape and the three other regions were analysed separately using the linear mixed-effects model and the coefficient of determination R^2 discussed in the previous section.

For the original landscape of the canton of Vaud, Figure 3 shows that sample size and modeling technique are the largest sources of variability in prediction accuracy, followed by the sampling design; the effects of the missing covariate and the dispersal process, i.e., the two possible sources of SAC, are less visible. Although there are differences among the species, all the graphs exhibit the same general pattern (see Supporting Information).

In order to quantify these visual impressions, we fitted the mixed-effects model (4) to the 32000 log(RMSE) values from our simulation. Inspection of the residuals indicated that the proposed mixed-effects model was appropriate for modeling the log(RMSE). The ANOVA is shown in Table 1. We then fitted the same model, excluding one factor at a time, and calculated the marginal and conditional R^2 values; see Table 2. The sample size n and the modeling technique are the most important factors. Varying n from 100 to 500 produces the greatest reduction in RMSE. The variation among modeling techniques can be huge, as seen in Figure 3, but this factor appears to be slightly less important than sample size in the ANOVA and in terms of reduction of the marginal R^2. There is a strong interaction between modeling technique and sample size, as can be seen in the ANOVA and in Figure 3. Although GAMs show generally poor performance in small samples, MaxEnt and RF are much less affected by small n. For the largest sample size, $n = 500$, GLMs are generally the best models, but for $n = 100$ they generally make poorer predictions than MaxEnt. Our results show that, in most cases, using MaxEnt on presence-absence data yields the best predictions in small samples. The effect of the sampling design is important, with a strong interaction with modeling technique: GLM and GAM are more sensitive to the choice of sampling design than are MaxEnt and RF (see the boxplots for the ten species in Supporting Information). The effects of the dispersal process and the missing covariate are the least important factors in our study.

For the three other regions of Switzerland, the ranking of the five factors changed only slightly. Sample size and modeling technique still had the largest impacts, while SAC due to the dispersal process remained the least important factor (see Supporting Information).

Discussion

In this paper we have proposed a simulation method for assessing the relative importance of factors in SDMs. The value of our general approach is that it provides a framework for simulation experiments, with a greater variety of species, landscapes and ranges of factors, that allow the assessment of the effects of SAC and further complications in other settings. In an application to five key factors in a real data framework (here plants in a mountain region), sample size and modeling technique had the largest relative effects on predictions, followed by the sampling design. In this illustration. the presence of SAC in the residuals, whether due to the dispersal process or a missing covariate, appeared to be of lower importance, despite being seen by many authors as a major issue (e.g., Segurado *et al.*, 2006; Dormann *et al.*, 2007). This suggests that the effect of SAC on prediction accuracy of SDMs may be relatively minor in studies similar to ours. Our analysis and virtual case study also confirmed the findings of Elith *et al.* (2006) and Wisz *et al.* (2008) regarding the performance of machine learning techniques such as MaxEnt, which generally outperform GLMs and GAMs in small samples. For the larger sample size, we found that GLMs were generally best, which is not surprising since our virtual species were generated using a GLM. The performance of MaxEnt in small samples is an interesting result that merits further investigation.

We chose values for factors, such as the sample size, to agree with applications related to our particular dataset, but of course these values are not universal. In particular, our simulation of SAC at the presence-absence level is intended to represent species dispersal at distances of around 1.5 kilometers. For the plant species of our dataset, this seems realistic (Vittoz & Engler, 2007). Using a higher correlation increases the relative importance of dispersal (see Supporting Information), but then seems less plausible. SAC resulting from an unobserved spatial covariate can yield dependence

at much larger distances, but can also be removed by including the corresponding predictor, so its importance is unclear. It can also be argued that this type of SAC does not affect the independence of the observations (Lichstein *et al.*, 2002; Diniz-Filho *et al.*, 2003; Guisan & Thuiller, 2005). Furthermore, we have only investigated the relative effects of the factors in terms of prediction accuracy, but other properties such as model selection and uncertainty estimation may be of interest. Although SAC here had relatively low impact on prediction accuracy, its effect on model selection and on the estimation of standard errors may be more important, since in a strong dispersal process the degrees of freedom may be lower than the number of observations.

The individual effects of different factors on SDM predictions have been the object of many previous studies, but to our knowledge, only Dormann *et al.* (2008), Diniz-Filho *et al.* (2009) and Garcia *et al.* (2012) considered the relative effects of factors on SDMs jointly. Diniz-Filho *et al.* (2009) and Garcia *et al.* (2012) evaluated the relative uncertainty due to the use of different modeling techniques and climate models to predict species responses to global change. Dormann *et al.* (2008) considered the relative effects of modeling technique, data uncertainty, collinearity correction and variable selection method, but did not include spatial autocorrelation. Furthermore, this latter study only used a single species and did not use as comprehensive a statistical framework as that used here. In contrast, our approach uses virtual data from a systematic experimental design that appropriately includes the three steps of species simulation, sampling procedure and modeling. Our framework is ecologically realistic and integrates randomness into the species distributions and sampling procedure. Other factors, such as location error, predictor error, multi-collinearity or the effect of using pseudo-absences, could also be assessed using the same framework.

Possibilities for future work include the study of the effects of niche complexity on prediction accuracy. Techniques such as MaxEnt or RF might be expected to be more efficient for modeling complex responses to predictors and thus might give more accurate predictions than GLMs or GAMs in realistic situations where species responses to predictors can be highly non-linear and depend on interactions between predictors. We acknowledge that the simulation of our virtual species may not be fully realistic and we encourage users of our methodological framework to explore alternative procedures for simulating virtual species. No universal method currently exists for simulating virtual species and more investigation is thus required. Another topic requiring further research is the relations between the different measures of accuracy used to assess the predictive power of SDMs. The RMSE appeared to be a natural choice in the framework of our simulations, but its relation to other metrics would benefit from further clarification.

We hope that our paper will pave the way toward more systematic analyses of factors affecting predictive models, especially when these and their predictions are subsequently to be used to derive scenarios of global change impact on biodiversity.

Acknowledgments: We thank M. G. Genton, P. J. Solomon, N. G. Yoccoz, R. B. O'Hara and anonymous reviewers for comments on the manuscript. This study was funded by the Swiss National Science Foundation in the context of the NCCR Plant Survival.

Data Accessibility:

- Species data: uploaded as online supporting information
- R scripts: uploaded as online supporting information

References

Albert, C.H., Yoccoz, N.G., Edwards, T.C., Graham, C.H., Zimmermann, N.E. & Thuiller, W. (2010) Sampling in ecology and evolution – bridging the gap between theory and practice. *Ecography*, **33**, 1028–1037.

Araújo, M.B. & Peterson, A.T. (2012) Uses and misuses of bioclimatic envelope modeling. *Ecology*, **93**, 1527–1539.

Baselga, A. & Araújo, M.B. (2009) Individualistic vs community modelling of species distributions under climate change. *Ecography*, **32**, 55–65.

Beale, C.M. & Lennon, J.J. (2012) Incorporating uncertainty in predictive species distribution modelling. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **367**, 247–258.

Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. & Elston, D.A. (2010) Regression analysis of spatial data. *Ecology Letters*, **13**, 246–264.

Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5–32.

Buisson, L., Thuiller, W., Casajus, N., Lek, S. & Grenouillet, G. (2010) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, **16**, 1145–1157.

Cablk, M., White, D. & Kiester, A.R. (2002) Assessment of spatial autocorrelation in empirical models in ecology. J. Scott, P. Heglund, M. Morrison, J. Haufler, M. Raphael, W. Wall & F. Samson, eds., *Predicting Species Occurrences: Issues of Accuracy and Scale*, pp. 429–440. Island Press, Covelo, CA.

Caruana, R. & Niculescu-Mizil, A. (2004) Data mining in metric space: an empirical analysis of supervised learning performance criteria. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '04, pp. 69–78. ACM, New York, NY, USA.

Chevan, A. & Sutherland, M. (1991) Hierarchical Partitioning. *The American Statistician*, **45**, 90–96.

Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A., Lodge, D.M., Pascual, M., Pielke, R., Pizer, W. & et al. (2001) Ecological forecasts: an emerging imperative. *Science*, **293**, 657–60.

Côté, I.M. & Reynolds, J.D. (2002) Conservation biology. Predictive ecology to the rescue? *Science*, **298**, 1181–1182.

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

Diniz-Filho, J.A.F., Mauricio Bini, L., Fernando Rangel, T., Loyola, R.D., Hof, C., Nogués-Bravo, D. & Araújo, M.B. (2009) Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. *Ecography*, **32**, 897–906.

Dormann, C.F. (2007) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, **16**, 129–138.

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

Dormann, C.F., Purschke, O., García Márquez, J.R., Lautenbach, S. & Schröder, B. (2008) Components of uncertainty in species distribution analysis: a case study of the Great Grey Shrike. *Ecology*, **89**, 3371–3386.

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M. & E. Zimmermann, N. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith, J. & Leathwick, J.R. (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology Evolution and Systematics*, **40**, 677–697.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

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.

Franklin, J. (2010) *Mapping Species Distributions. Spatial Inference and Prediction*. Cambridge University Press, Cambridge, UK.

Garcia, R.A., Burgess, N.D., Cabeza, M., Rahbek, C. & Araújo, M.B. (2012) Exploring consensus in 21st century projections of climatically suitable areas for African vertebrates. *Global Change Biology*, **18**, 1253–1269.

Graham, C.H., Ferrier, S., Huettman, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology & Evolution*, **19**, 497–503.

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

Hastie, T. (2013) *gam: Generalized Additive Models*. R package version 1.08.

Hastie, T. & Fithian, W. (2013) Inference from presence-only data; the ongoing controversy. *Ecography*, **36**, 864–867.

Hastie, T.J. & Tibshirani, R.J. (1990) *Generalized Additive Models*. London: Chapman & Hall.

Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2013) *dismo: Species distribution modeling*. R package version 0.8-17.

Hirzel, A. & Guisan, A. (2002) Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling*, **157**, 331–341.

Liaw, A. & Wiener, M. (2002) Classification and Regression by randomForest. *R News*, **2**, 18–22.

Lichstein, J.W., Simons, T.R., Shriner, S.A. & Franzreb, K.E. (2002) Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, **72**, 445–463.

Liu, C., White, M. & Newell, G. (2011) Measuring and comparing the accuracy of species distribution models with presence–absence data. *Ecography*, **34**, 232–243.

Mac Nally, R. (2000) Regression and model-building in conservation biology, biogeography and ecology: The distinction between –and reconciliation of –'predictive'and 'explanatory'models. *Biodiversity & Conservation*, **9**, 655–671.

Mason, S.J. & Graham, N.E. (2002) Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society*, **128**, 2145–2166.

McCullagh, P. & Nelder, J.A. (1989) *Generalized Linear Models (Second Edition)*. London: Chapman & Hall.

Merow, C., Smith, M.J. & Silander, J.A. (2013) A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography*, pp. no–no.

Miller, J.A. (2014) Virtual species distribution models: Using simulated data to evaluate aspects of model performance. *Progress in Physical Geography*, **38**, 117–128.

Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133–142.

Pereira, H., Leadley, P., Proenca, V., Alkemade, R., Scharlemann, J., Fernandez-Manjarres, J., Araujo, M., Balvanera, P., Biggs, R., Cheung, W., Chini, L., Cooper, H., Gilman, E., Guenette, S., Hurtt, G., Huntington, H., Mace, G., Oberdorff, T., Revenga, C., Rodrigues, P., Scholes, R., Sumaila, U. & Walpole, M. (2010) Scenarios for global biodiversity in the 21st century. *Science*, **330**, 1496–1501.

Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martínez-Meyer, E., Nakamura, M. & Araújo, M.B. (2011) *Ecological Niches and Geographic Distributions*. Princeton University Press, Princeton, USA.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Development Core Team (2013) *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-109.

Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C.F., Vittoz, P. & Guisan, A. (2013) The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Global Ecology and Biogeography*, **22**, 52–63.

R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

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

Renner, I.W. & Warton, D.I. (2013) Equivalence of MAXENT and Poisson Point Process Models for Species Distribution Modeling in Ecology. *Biometrics*, **69**, 274–281.

Schröter, D., Cramer, W., Leemans, R., Prentice, I.C., Araújo, M.B., Arnell, N.W., Bondeau, A., Bugmann, H., Carter, T.R., Gracia, C.A. & et al. (2005) Ecosystem service supply and vulnerability to global change in Europe. *Science*, **310**, 1333–1337.

Segurado, P., Araújo, M.B. & Kunin, W.E. (2006) Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, **43**, 433–444.

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

Tate, R.F. (1954) Correlation Between a Discrete and a Continuous Variable. Point-Biserial Correlation. *The Annals of Mathematical Statistics*, **25**, 603–607.

Thuiller, W., Araújo, M., Pearson, R., Whittaker, R., Brotons, L. & Lavorel, S. (2004) Uncertainty in predictions extinction risk. *Nature*, **430**, 34.

Venables, W.N. & Ripley, B.D. (2002) *Modern Applied Statistics with S. Fourth Edition*. Springer. ISBN 0-387-95457-0.

Vittoz, P. & Engler, R. (2007) Seed dispersal distances: a typology based on dispersal modes and plant traits. *Botanica Helvetica*, **117**, 109–124.

Wisz, M.S. & Guisan, A. (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, **9**, 8.

Wisz, M.S., Hijmans, R.J., Peterson, A.T., Graham, C.H., Guisan, A. & NCEAS Predicting Species Distributions Working Group (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H. & Veran, S. (2013) Presence-only modelling using MAXENT: when can we trust the inferences? *Methods in Ecology and Evolution*, **4**, 236–243.

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

Zurell, D., Berger, U., Cabral, J.S., Jeltsch, F., Meynard, C.N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B. & Grimm, V. (2010) The virtual ecologist approach: simulating data and observers. *Oikos*, **119**, 622–635.

Table 1: Analysis of variance for the linear mixed-effects model (4). Only factors for which the corresponding p-value is smaller than 0.05 are shown. A colon denotes interaction.

Error level	Factor	Df	Sum Sq	Mean Sq	F-value	$\log_{10}(p)$
Species	residual	9	124.686	13.854		
Simulation in species	missing	1	15.835	15.835	159.658	−30.177
	dispersal	1	12.363	12.363	124.647	−24.568
	residual	389	38.582	0.099		
Sampling in simulation	n	1	1723.328	1723.328	9486.835	−1009.041
	design	1	373.194	373.194	2054.412	−354.34
	missing:n	1	15.62	15.62	85.985	−19.516
	dispersal:n	1	4.69	4.69	25.818	−6.404
	missing:design	1	2.449	2.449	13.481	−3.612
	missing:n:design	1	1.618	1.618	8.909	−2.544
	residual	3586	651.414	0.182		
Within sampling	technique	3	480.571	160.19	2313.483	−1186.59
	missing:technique	3	7.964	2.655	38.338	−23.922
	dispersal:technique	3	3.529	1.176	16.987	−10.28
	n:technique	3	702.992	234.331	3384.229	−1594.082
	design:technique	3	60.178	20.059	289.699	−180.819
	missing:n:technique	3	11.517	3.839	55.442	−34.856
	dispersal:n:technique	3	1.003	0.334	4.826	−2.633
	missing:design:technique	3	0.929	0.31	4.471	−2.416
	n:design:technique	3	1.346	0.449	6.479	−3.65
	missing:n:design:technique	3	1.662	0.554	8	−4.598
	residual	11952	827.58	0.069		

Table 2: Marginal and conditional R^2 (Nakagawa & Schielzeth, 2013) for the full model (4) and the five sub-models with one factor, and all its interactions with the other factors, excluded at a time.

Model	full model	−missing	−dispersal	−n	−design	−technique
Marginal R^2	0.674	0.662	0.669	0.188	0.587	0.423
Conditional R^2	0.782	0.777	0.781	0.594	0.766	0.450

Figure 1: Simulation of the first virtual species. We use the five real predictors in (a) and the probit function to define the map of probability of occurrence in (b). The presence-absence process is generated with SAC in (c) using a Gaussian random field having an effective range of 1.5 km, and in (d) without SAC using independent normal variables. For each configuration of dispersal (true or false), a total of 100 presences (red dots) and absences (blue dots) are sampled using a simple random design in (e) and (g), and a road-based design in (f) and (h).

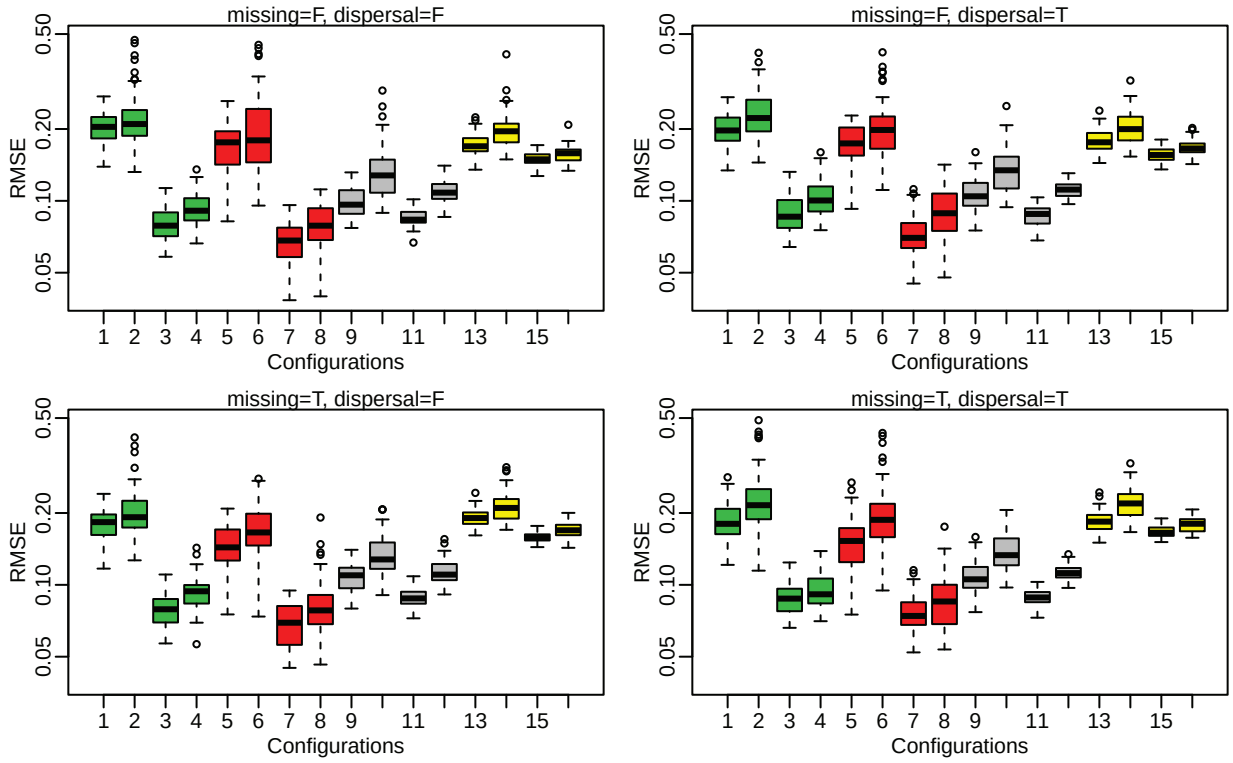


Figure 2: Boxplots of RMSE for the first simulated species, with a log scale on the vertical axis. Each box shows the variation due to the 5 samples for each of the 10 simulations and for one particular configuration of factors (missing covariate, dispersal, modeling technique, sample size n and sampling design), and thus contains 50 RMSE values, each equal to the root mean squared difference of the estimated and the true probabilities of presence over the 5000 sites of the test samples. Each of the four panels corresponds to a different combination of the factors *missing* and *dispersal*. In each panel the four modeling techniques are distinguished by colors: generalized additive models (GAMs) in green, generalized linear models (GLMs) in red, maximum entropy (MaxEnt) in grey, and random forests (RF) in yellow. Inside the color groups, boxplots are first separated by the sample size n (100 or 500, to the left or right), and within sample size by the sampling design (simple random or road-based, to the left or right). Thus, configurations 1–4 in the top left panel correspond to the 50 RMSE values for GAMs with (sample size, sampling design) settings (100, simple random), (100, road-based), (500, simple random) and (500, road-based), respectively, for data with no missing covariate and no dispersal process. Configurations 5–8 in the same panel show the RMSEs for GLMs, configurations 9–12 for MaxEnt, and 13-16 for RF, all with the corresponding settings.

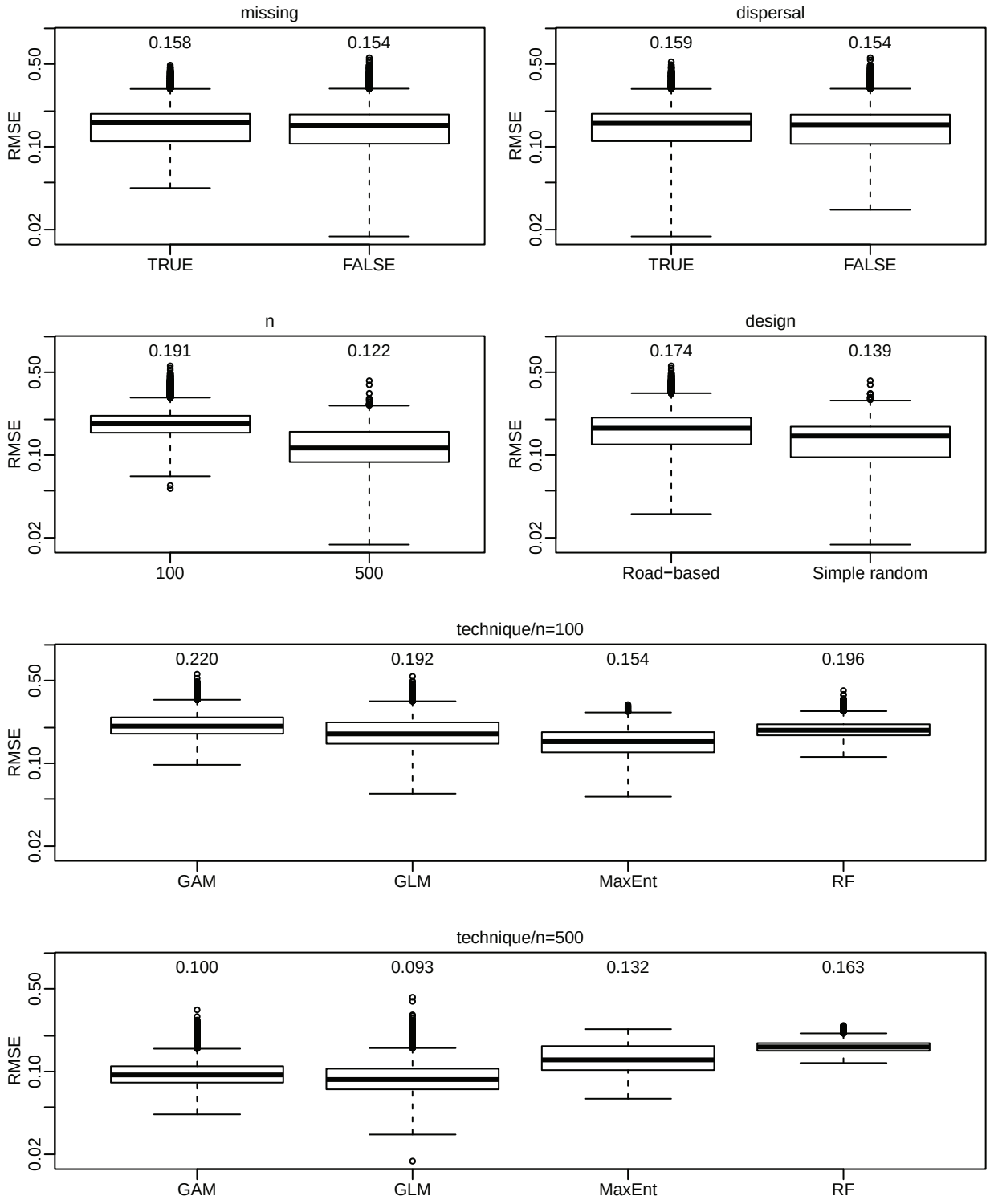


Figure 3: Boxplots of all RMSE values (for the ten virtual species) for each configuration of the five factors. The mean of the RMSE values for each configuration is printed over the corresponding boxplot.

Supporting Information for "Measuring the relative effect of factors affecting species distribution model predictions"

Emeric Thibaud,[1*] Blaise Petitpierre,[2] Olivier Broennimann,[2]
Anthony C. Davison,[1] and Antoine Guisan[2,3]

[1]Chair of Statistics, Ecole Polytechnique Fédérale de Lausanne,
EPFL-FSB-MATHAA-STAT, Station 8, 1015 Lausanne, Switzerland
[2]Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland
[3]Faculty of Geoscience and the Environment, University of Lausanne, 1015 Lausanne, Switzerland

[*]To whom correspondence should be addressed; E-mail: emeric.thibaud@epfl.ch.

Contents

S1 Data

Our simulation uses a real Alpine landscape in the western Swiss Alps ($7°2'$–$7°14'$ E; $46°28'$–$46°31'$ N). This landscape represents an area of about 700 km^2 of the canton of Vaud, comprising $1,127,599$ pixels at a resolution size of 25 meters. It has been intensively sampled (Pottier *et al.*, 2013), and this helps in providing realistic estimates for parameters for the simulation. We used five real climatic and topographic predictors known for their importance in shaping plant distributions (Zimmermann & Kienast, 1999; Randin *et al.*, 2006), labelled $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_5$:

- the annual sum of degree days above three degrees (*ddeg300*);
- the annual average number of frost days during the growing season (*sfroyy*) from existing maps (Zimmermann & Kienast, 1999);
- a moisture index between June and August (*mind68*);
- daily average global potential shortwave radiation per month (*sumradyy*); and
- topographic position (*topos*, Randin *et al.*, 2006).

We simulated our virtual species based on real species data. The species dataset comprises presence-absence records at 912 locations for 260 plant species. Our simulation was based on parameters from ten representative species in the study area (Table S1). These are herbs, growing from the lowland to the alpine levels with different ecological tolerances. No trees or shrubs were included because forestry and pasturing may have truncated their distribution, especially at higher altitudes. These species were used as realistic references for the α and β parameters of the probit models used to build our virtual species distributions, and were also used to estimate a realistic spatial autocorrelation coefficient; see §S3.

S2 The root mean squared error

Many measures of model accuracy (Fielding & Bell, 1997; Caruana & Niculescu-Mizil, 2004; Liu *et al.*, 2011) can be used in SDMs. Because in practice the true probabilities of presence p_i are unknown, accuracy measures used in SDMs compare the observed presences and absences in the test sample with the predicted probabilities $\hat{p}_i$. Two popular measures are the area under the receiver operating characteristic curve (AUC, Mason & Graham, 2002) and the point-biserial correlation (COR, Tate, 1954).

For our simulation, the true distribution of the virtual species is known, which suggest comparing the p_i with the predictions $\hat{p}_i$. We used the root mean squared error (RMSE, Caruana & Niculescu-Mizil, 2004; Liu *et al.*, 2011):

$$\text{RMSE} = \sqrt{\frac{1}{n'} \sum_{i \in \mathcal{T}} (\hat{p}_i - p_i)^2}.$$

The RMSE has a simple interpretation as the mean distance between predicted and true probabilities over the locations of the test sample. The decomposition of the expectation of its square as

$$\text{E}\left\{ \frac{1}{n'} \sum_{i \in \mathcal{T}} (\hat{p}_i - p_i)^2 \right\} = \frac{1}{n'} \sum_{i \in \mathcal{T}} \left[\text{var}(\hat{p}_i) + \{p_i - \text{E}(\hat{p}_i)\}^2 \right],$$

shows that it penalizes both the variability and the bias of the estimator.

Because the RMSE uses only the true p_i, it corresponds to the accuracy related to the estimation of the fundamental niche of the species, excluding possible undesirable effects of the SAC at the presence-absence level. Use of independent test samples, i.e., pixels that are not in the training samples, avoids bias in estimation of prediction accuracy, and because the RMSE compares predictions with true presence probabilities, and not observed presences or absences, it is unaffected by any proximity of the pixels used for the training and test samples. In practice, when using cross-validation and the AUC or COR, modelers must use well-separated pixels for the training and test samples in order to reduce any possible effects of spatial autocorrelation between the training and test samples.

The AUC and COR can be calculated in our simulation framework by generating presences and absences at each location of the test sample based on the true p_i. Two approaches to generate these presences and absences are possible. One approach generates correlated presences and absences using correlated normal variables that mimic the SAC present in the training sample. To avoid bias in the estimation of the RMSE these presences and absences must be independent of those in the training sample. This approach corresponds to the estimation on a test sample independent of the training sample but within which the records are dependent because of SAC. A second approach uses independent normal variables to generate independent presences and absences without SAC. The degrees of freedom of the test sample for the first approach are smaller than for the second

approach and the variability of the RMSE is larger. Compared to the RMSE which uses the true p_i, measures based on simulated presences and absences are more variable.

In our application, the distribution of log(RMSE) was found to be approximately normal, which was appropriate for the use of the linear mixed-effect model and the analysis of variance. We also calculated the AUC and COR but their distributions makes further analysis more complicated. The use of measures other than log(RMSE) within our framework seems quite feasible, at least in principle, but requires further investigation.

S3 Estimating spatial autocorrelation

Here we discuss the choice of the strength of spatial autocorrelation (SAC) corresponding to a dispersal process. To determine a sensible value for the range parameter λ chosen in our simulation (see main article), we estimated the range of spatial autocorrelation on the real species we used (see Section S1) by fitting a spatial probit model similar to those used to simulate the virtual species.

Given predictors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_5$ we let $\boldsymbol{X}$ denote the $N \times 5$ matrix of predictors, in which each row corresponds to a different location. We use the probit function to relate the predictors to the presence probabilities, i.e.,

$$p(\boldsymbol{X}) = \Phi\{\eta(\boldsymbol{X})\}, \tag{S1}$$

where Φ is the standard Gaussian cumulative distribution function. We include quadratic terms for each predictor and we do not consider interactions, so that

$$\eta(\boldsymbol{X}) = \alpha_0 + \sum_{j=1}^{5} \alpha_j (\boldsymbol{x}_j - \beta_j)^2$$

with coefficients α_j and β_j to be estimated. The standard GLM probit model assumes independence of the outcomes (in our case, presences–absences) but here we consider a spatial model in which the marginal distributions are unchanged but presences and absences tend to cluster.

In the standard probit model, presences and absences are obtained by generating independent standard normal variables $\gamma_1, \ldots, \gamma_N$ for the N locations, writing η_i for the linear predictor at the ith location, and defining the binary responses at those locations by

$$Y_i = \begin{cases} 1, & \gamma_i \leq \eta_i, \\ 0, & \gamma_i > \eta_i. \end{cases} \tag{S2}$$

The spatial probit model considered here presupposes that presences and absences are obtained using correlated γ_i in (S2): we replace the independent γ_i by a Gaussian process $\gamma(\boldsymbol{s})$ with zero mean, unit variance and correlation function ρ, where $\boldsymbol{s}$ is the location of the record. The range of SAC can be precisely controlled, as the parameters of the correlation function ρ control the range of dependence. The marginal probability of presence at location $\boldsymbol{s}$ is still defined from Equation (S1) but now the presences–absences tend to cluster. Pairwise marginal distributions are given by

$$\Pr(Y_i = y_i, Y_j = y_j) = \begin{cases} \Phi_2(\eta_i, \eta_j; \rho_{ij}), & y_i = 1, y_j = 1, \\ \Phi(\eta_i) - \Phi_2(\eta_i, \eta_j; \rho_{ij}), & y_i = 1, y_j = 0, \\ \Phi(\eta_j) - \Phi_2(\eta_i, \eta_j; \rho_{ij}), & y_i = 0, y_j = 1, \\ 1 - \Phi(\eta_i) - \Phi(\eta_j) + \Phi_2(\eta_i, \eta_j; \rho_{ij}), & y_i = 0, y_j = 0, \end{cases} \tag{S3}$$

where $\Phi_2(\ ,\ ; \rho)$ denotes the bivariate standard Gaussian cumulative distribution function with correlation ρ. Here we use $\rho_{ij} = \exp(-h_{ij}/\lambda)$, where h_{ij} is the distance (in km) between locations $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$ and λ is a range parameter to be chosen.

We use a composite likelihood approach to estimate the parameters of the spatial probit model. Pairwise distributions (S3) are used to construct the pairwise log-likelihood (Varin *et al.*, 2011) function

$$\ell(\boldsymbol{\psi}) = \sum_{1 \leq i < j \leq N} w_{ij} \log \Pr(Y_i = y_i, Y_j = y_j), \tag{S4}$$

where the sum is over all distinct pairs (i, j), the w_{ij} are weights to be chosen, and $\boldsymbol{\psi}$ is the vector of model parameters. Properties of composite likelihood estimation are discussed by Varin *et al.* (2011). In particular, under similar assumptions to that of classical likelihood estimation, it can be show that the maximum composite likelihood $\hat{\boldsymbol{\psi}} = \operatorname{argmax}_{\boldsymbol{\psi}} \ell(\boldsymbol{\psi})$ estimator is consistent and that its distribution is asymptotically normal, with mean $\boldsymbol{\psi}$ and covariance matrix of the standard "sandwich" form (Varin *et al.*, 2011). Using composite likelihood estimation instead of the full likelihood usually results in a loss in efficiency. No closed form expression exist for $\hat{\boldsymbol{\psi}}$ but the pairwise log-likelihood function $\ell(\boldsymbol{\psi})$ can be maximized using the R function optim. Because the number of pairs can be large, maximizing this likelihood can be burdensome, so we used two shortcuts to accelerate the optimization. First, we estimated the regression parameters in the η_i by using a probit model that wrongly assumes that the responses are all independent (i.e., a GLM with a probit link). Such estimators are consistent, but their variances are mis-estimated. The estimates of η_i are then plugged into the likelihood (S4), so that then only only the dependence parameter λ need be estimated. Second, to reduce the number of terms involved, we excluded some pairs, by setting $w_{ij} = 0$ for pixels more than 2.4 km apart.

We use the spatial probit model and pairwise likelihood estimates to find estimates of λ for each species. The results are shown in Table S1. These estimates are sensitive to the choice of the subset of pairs used in the likelihood, so we chose to fix $\lambda = 0.5$ in our simulation, though the estimated values of λ are all smaller than 0.39. Moreover, choosing $\lambda = 0.5$ implies correlation lower than 0.05 at distances over 1.5 km. Similar species dispersal values have been found by Vittoz & Engler (2007).

Table S1: Names of the real species used for the simulation and estimated values of spatial autocorrelation using a spatial probit model.

	Species name	$\hat{\lambda}$	Effective range (km)
1	Festuca pratensis sl.	0.13	0.39
2	Prunella vulgaris	0.27	0.79
3	Veronica chamaedrys	0.19	0.57
4	Taraxacum officinale aggr.	0.03	0.08
5	Plantago lanceolata	0.39	1.18
6	Cerastium fontanum sl.	0.02	0.05
7	Agrostis capillaris	0.26	0.79
8	Alchemilla xanthochlora aggr.	0.30	0.90
9	Leontodon hispidus sl.	0.23	0.68
10	Festuca rubra aggr.	0.29	0.88

S4 Varying the strength of spatial autocorrelation

We investigate how the conclusions of the main paper on the relative importance of the five factors change when varying the strength of SAC in the dispersal process. We choose different values of the range parameter λ of the correlation function of the underlying Gaussian process used to mimic a dispersal process, representing weaker or stronger SAC at the presence–absence simulation level. We then calculate the marginal R^2 (Nakagawa & Schielzeth, 2013) for the full model and each sub-model to determine the relative importance of the factors. The marginal R^2 values obtained are shown in Table S2. For a Gaussian process having an effective range of 5 km, the effect of dispersal is comparable to that of sampling design, for 10 km it is comparable to that of modeling technique and for 15 km it is comparable to that of sample size n. Not surprisingly, increasing the strength of SAC for the dispersal process increases its relative effect on prediction accuracy, and it becomes one of the most important factors for large SAC, but such large dispersal values are unrealistic for the species considered in our study, see Table S1 and Vittoz & Engler (2007).

Table S2: Marginal R^2 for the full model (3) and the five sub-models with one factor (and all its interactions with the other factors) excluded at a time, and for various strengths of spatial autocorrelation for the dispersal process. Different values for the range parameter λ correspond to different effective ranges (distance after which the correlation drops below 5%).

Effective range (km)	λ	full model	−missing	−dispersal	−n	−design	−technique
1.5	0.5	0.674	0.662	0.669	0.188	0.587	0.423
5	1.67	0.644	0.632	0.566	0.233	0.559	0.412
10	3.34	0.651	0.641	0.446	0.315	0.579	0.428
15	5	0.670	0.661	0.372	0.381	0.612	0.451

S5 External validation

We investigated the ranking of the factors in terms of the accuracy of model predictions estimated on different landscapes. We measured the accuracy of predictions obtained by the different combinations of the factors using the RMSE calculated by predicting species distributions in other regions where different correlations among the predictors occur. In the main paper, we used the landscape VD of the Vaud Alps and ten real species present in this region. For external validation we use three other regions of Switzerland where the same species are present but that present different topo-climatic conditions: the lower Engadine (EN), the canton of Neuchâtel (NE), and the south part of the canton of Ticino (TI). Table S3 summarises the main characteristics of the four regions. We use the same climatic and topographic predictors as for VD, with the same pixel resolution of 25 meters.

Table S3: Coordinates, number of pixels, and mean values for the predictors for the four regions of Switzerland used in our simulation (VD, EN, NE and TI).

	Pixels	Coordinates	ddeg300	sfroyy	mind68	sumradyy	topos
VD	$1,127,599$	$7°2'$–$7°14'$ E; $46°28'$–$46°31'$ N	1788	21	177	199866	-6.2
EN	$1,597,016$	$9°56'$–$10°29'$ E; $46°36'$–$47°0'$ N	735	204	244	194969	-3.8
NE	$1,146,942$	$6°25'$–$7°5'$ E; $46°50'$–$47°9'$ N	1861	15	-176	217983	0.8
TI	$874,996$	$8°45'$–$9°9'$ E; $45°49'$–$46°11'$ N	2693	4	233	199602	-0.8

The theoretical probability maps of presence for the ten virtual species in the landscapes EN, NE and TI are obtained via the probit models used to generate the virtual species in VD. Figures S10–S19 show the maps of presence probabilities for the ten virtual species in the four landscapes.

Predictions from modeling techniques are obtained using exactly the same procedure as for internal validation. One RMSE value is computed by comparing predicted probabilities of presence with the truth over 5000 locations of a test sample.

The results of the fit of mixed-effects models for each validation landscape are shown in Table S4. Using the R^2 for the full models and all sub-models with one factor excluded at a time, we rank the importance of the factors for each landscape. For VD, NE and TI, the sample size is the most important factor followed by the modeling technique. For EN, the effect of modeling technique is larger than the effect of the sample size. For all landscapes except TI, the effect of the sampling design is larger than that of the missing covariate. Dispersal is the least important factor for all the landscapes, though it ties with sampling design in TI.

Correlation matrices between the five predictors *ddeg300*, *sfroyy*, *mind68*, *sumradyy* and *topos*

(in this order) for the four landscapes are:

$$\mathrm{cor}_{\mathrm{VD}} = \begin{pmatrix} 1.00 & -0.49 & -0.85 & 0.19 & -0.27 \\ -0.49 & 1.00 & 0.55 & -0.18 & 0.23 \\ -0.85 & 0.55 & 1.00 & -0.54 & 0.21 \\ 0.19 & -0.18 & -0.54 & 1.00 & 0.09 \\ -0.27 & 0.23 & 0.21 & 0.09 & 1.00 \end{pmatrix},$$

$$\mathrm{cor}_{\mathrm{EN}} = \begin{pmatrix} 1.00 & -0.79 & -0.70 & 0.00 & -0.38 \\ -0.79 & 1.00 & 0.77 & 0.04 & 0.47 \\ -0.70 & 0.77 & 1.00 & -0.27 & 0.36 \\ 0.00 & 0.04 & -0.27 & 1.00 & 0.11 \\ -0.38 & 0.47 & 0.36 & 0.11 & 1.00 \end{pmatrix},$$

$$\mathrm{cor}_{\mathrm{NE}} = \begin{pmatrix} 1.00 & -0.67 & -0.81 & 0.10 & -0.28 \\ -0.67 & 1.00 & 0.49 & -0.08 & 0.07 \\ -0.81 & 0.49 & 1.00 & -0.43 & 0.36 \\ 0.10 & -0.08 & -0.43 & 1.00 & 0.08 \\ -0.28 & 0.07 & 0.36 & 0.08 & 1.00 \end{pmatrix},$$

$$\mathrm{cor}_{\mathrm{TI}} = \begin{pmatrix} 1.00 & -0.62 & -0.58 & 0.11 & -0.32 \\ -0.62 & 1.00 & 0.37 & 0.01 & 0.30 \\ -0.58 & 0.37 & 1.00 & -0.70 & 0.05 \\ 0.11 & 0.01 & -0.70 & 1.00 & 0.11 \\ -0.32 & 0.30 & 0.05 & 0.11 & 1.00 \end{pmatrix}.$$

Table S4: Marginal R^2 for the full model (3) and the five sub-models with one factor (and all its interactions with the other factors) excluded at a time, and for various external validation landscapes.

Landscape	full model	−missing	−dispersal	−n	−design	−technique
VD (internal)	0.674	0.662	0.669	0.188	0.587	0.423
EN	0.250	0.247	0.250	0.162	0.166	0.146
NE	0.434	0.415	0.428	0.085	0.403	0.265
TI	0.325	0.248	0.319	0.082	0.319	0.223

S6 R scripts and data

We provide the R scripts and data to run all analyses presented in the article. See the separate file entitled "Scripts_Thibaud_etal_MEE.zip", which contains:

1. main.R: The main R script for running the simulation, plotting the results and calculating the coefficients R^2.
2. functionsFI.R: Some functions for the simulation.
3. create_species.R: R scripts for creating the virtual species.
4. species_VD.R: R scripts for calculating the spatial autocorrelation for VD species.
5. fit_spatialprobit.R: Function to fit the spatial probit model described in §S3.
6. save_matrixresults.R: Function to save the results of the simulation (RMSE) and the corresponding configurations in a matrix.
7. plot_results.R: Function to plot the results.
8. anova.R: Function to do the ANOVA for the log(RMSE) values.
9. factor_importance_R2.R: Function to calculate the coefficients R^2.
10. extern_valid.R: Fonction for validation on other landscapes.
11. pred_VD.Rdata, pred_EN.Rdata, pred_NE.Rdata, pred_TI.Rdata: Data file for the real predictors in VD, EN, NE and TI.
12. distroad_VD.Rdata: Data file for the sampling design (distance to the nearest-road).
13. datasp_VD.Rdata: Data file for the 10 real species used in VD.
14. FI-VD.Rdata, FI-VD-extEN.Rdata, FI-VD-extNE.Rdata, FI-VD-extTI.Rdata: The results for the simulations of the paper.

To run the simulation process, only the file main.R need be open. It must be run line by line and it will load the data and source the necessary other files. The script species_VD.R is independent of the simulation and allows estimation of the spatial autocorrelation for the real species using a spatial probit model. Finally, the data files "FI-VD.Rdata", "FI-VD-extEN.Rdata", "FI-VD-extNE.Rdata" and "FI-VD-extTI.Rdata" contain the results of the simulations presented in the paper.

S7 Supporting Table and Figures

Additional Tables and Figures:

Table S5: Some references treating the effects of modeling factors on SDM predictions.

Figures S1–S9: Boxplots of RMSE values for the simulated species 2–10 and for each combination of the factors we considered. Validation is performed on the same landscape (VD).

Figures S10–S19: Maps of presence probability for the simulated species 1–10 in the landscapes VD, EN, NE and TI.

Table S5: References treating the effects of modeling factors on SDM predictions.

Modeling factor	References
Modeling techniques	Hirzel *et al.* (2001);Moisen & Frescino (2002);Thuiller (2003);Segurado & Araújo (2004); Elith *et al.* (2006);Guisan *et al.* (2007a);Meynard & Quinn (2007)*;Tsoar *et al.* (2007); Elith & Graham (2009)*
Grain	Hortal *et al.* (2006);McPherson *et al.* (2006);Guisan *et al.* (2007b)
Location error	Graham *et al.* (2008);Johnson & Gillingham (2008);Naimi *et al.* (2011)*
Pseudo-absence selection	Zaniewski *et al.* (2002);Phillips *et al.* (2009);Wisz & Guisan (2009)*;Vanderwal *et al.* (2009)
Sample size	Hirzel *et al.* (2002);Stockwell & Peterson (2002);Kadmon *et al.* (2003);Reese *et al.* (2005)*; Hernandez *et al.* (2006);Wisz *et al.* (2008);Jiménez-Valverde *et al.* (2009)*
Multi-collinearity	Graham (2003)†
Geographic extent	Thuiller *et al.* (2004)
Threshold criteria for binarizing predictions	Freeman & Moisen (2008);Liu *et al.* (2005)
Data completeness	Peterson & Cohoon (1999);Kadmon *et al.* (2003)
Data characteristics	Kadmon *et al.* (2003);Guisan *et al.* (2007a)
Species characteristics	McPherson *et al.* (2004)†;Guisan *et al.* (2007a);M McPherson & Jetz (2007); Broennimann *et al.* (2012)†
Sampling design	Austin & Adomeit (1991)*;Hirzel *et al.* (2002);Kadmon *et al.* (2003);Kadmon *et al.* (2004); Reese *et al.* (2005)*;Phillips *et al.* (2009);Albert *et al.* (2010)*
Spatial autocorrelation	Cablk *et al.* (2002);Segurado *et al.* (2006)†;Dormann (2007);Dormann *et al.* (2007)*

*use simulated data

† use simulated and real data

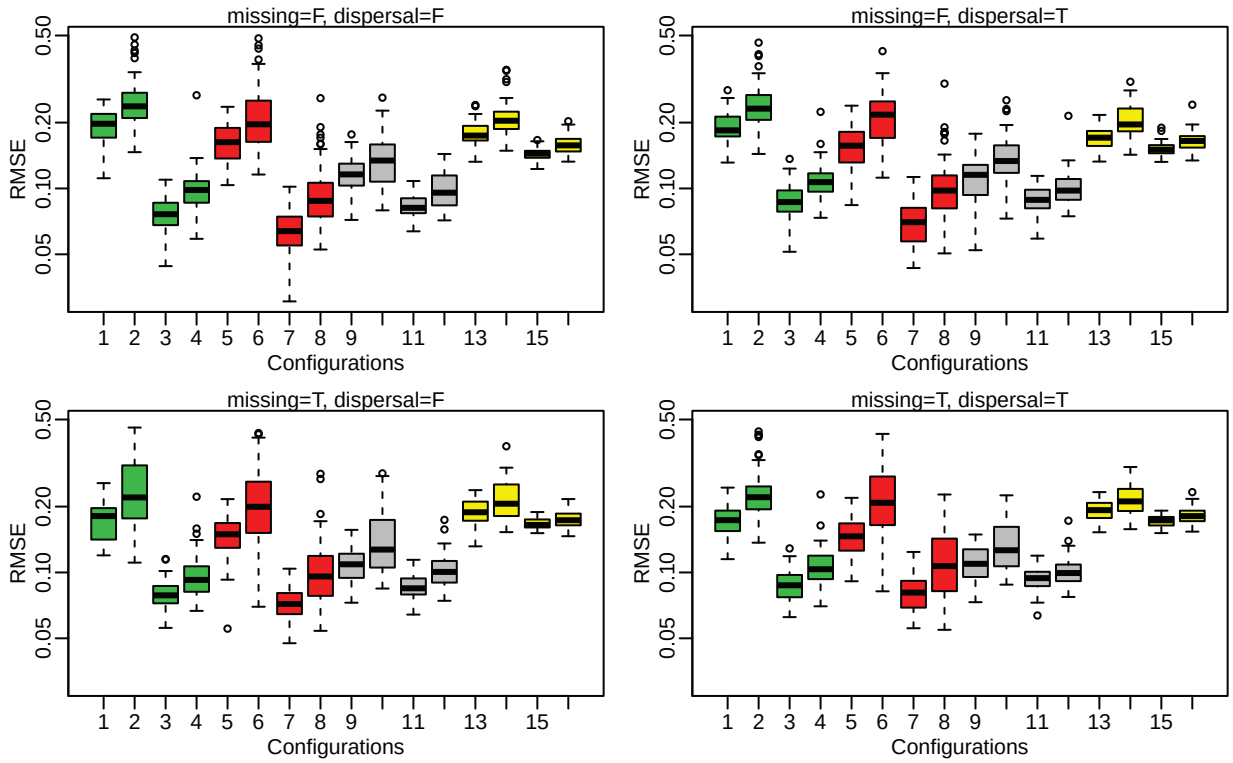


Figure S1: Boxplots of RMSE for simulated species 2, on a log scale. Each boxplot shows the variation due to the 5 samples for each of the 10 simulations and for one particular configuration of factors (missing covariate, dispersal, modeling technique, sample size n and sampling design), and thus contains 50 RMSE values, each equal to the root mean squared difference of the estimated and the true probabilities of presence over the 5000 sites of the test samples. Each panel corresponds to a different configuration of the factors *missing* and *dispersal*. In each panel colors correspond to the different modeling techniques: GAMs in green, GLMs in red, MexEnt in grey, and RF in yellow. Inside the color groups, boxplots are first separated by the sample size n (100 or 500, to the left or right) and within sample size by the sampling design (simple random or road-based, to the left or right). Thus, configurations 1–4 correspond to results for GAMs with (sample size, sampling design) settings (100, simple random), (100, road-based), (500, simple random) and (500, road-based), respectively.

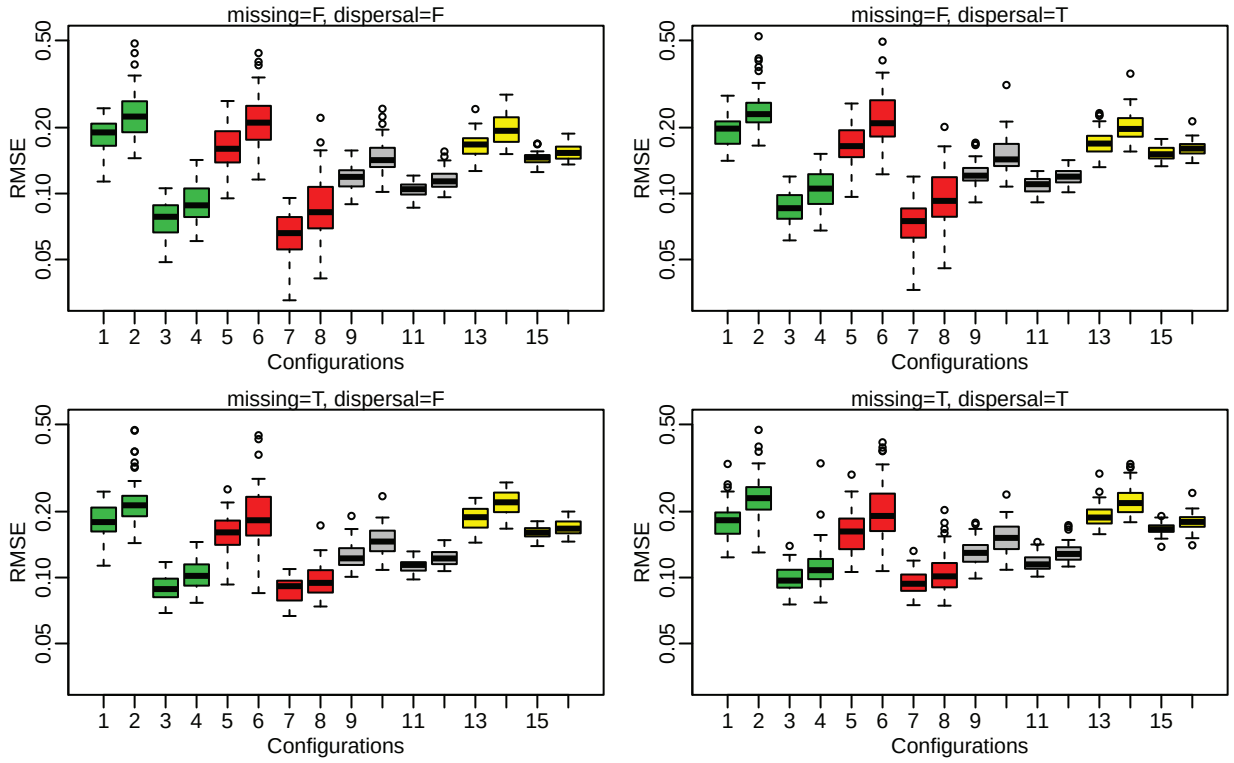


Figure S2: Boxplots of RMSE for simulated species 3. Same caption as Figure S1.

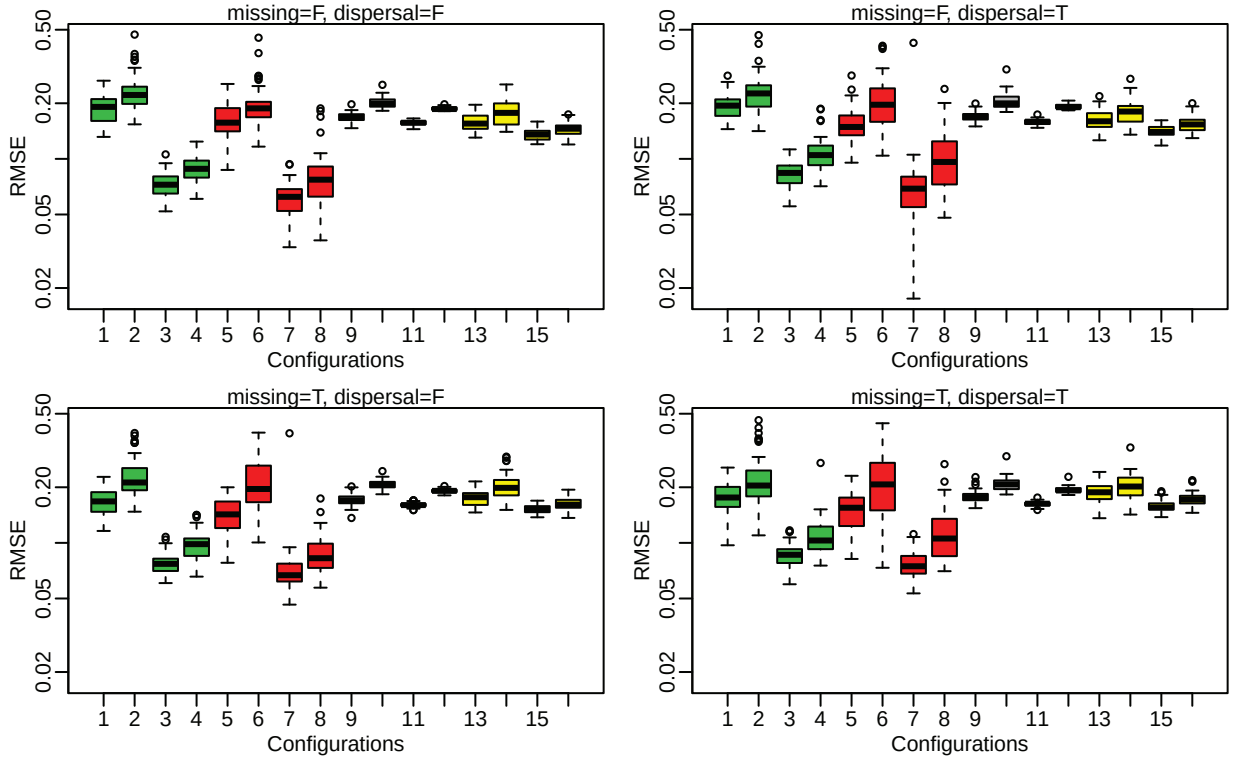


Figure S3: Boxplots of RMSE for simulated species 4. Same caption as Figure S1.

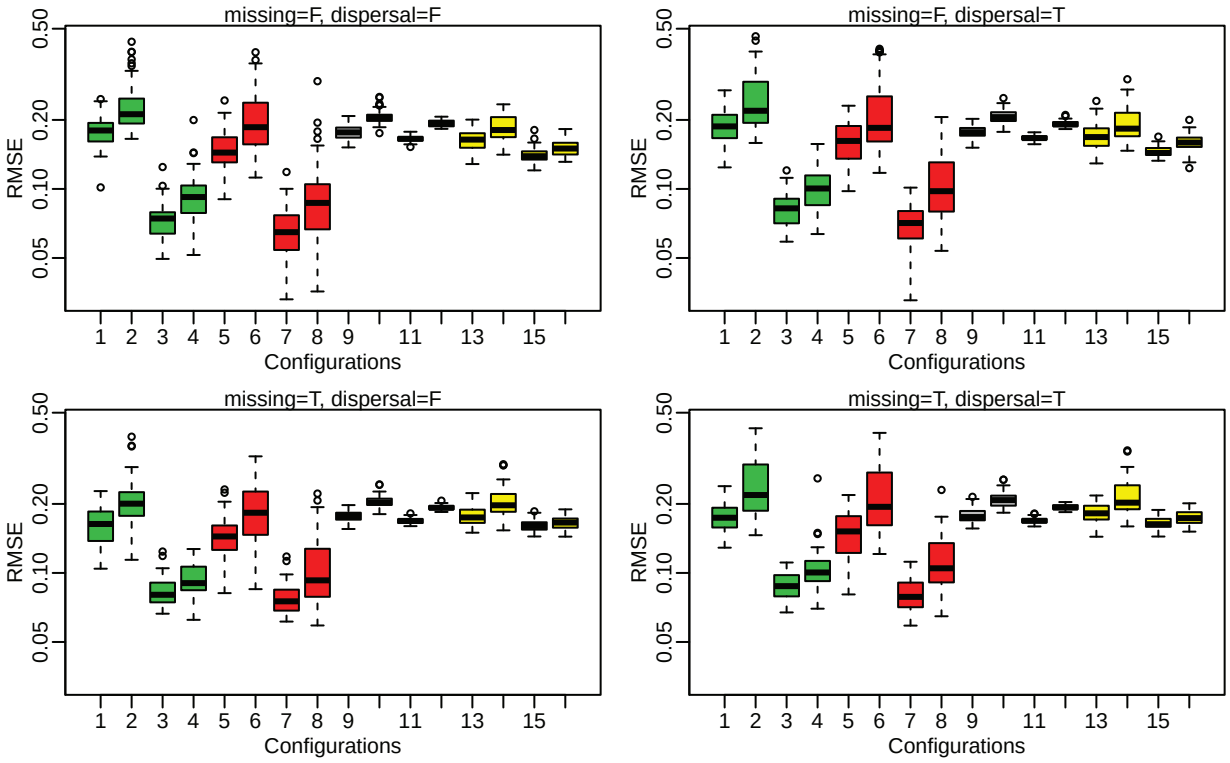


Figure S4: Boxplots of RMSE for simulated species 5. Same caption as Figure S1.

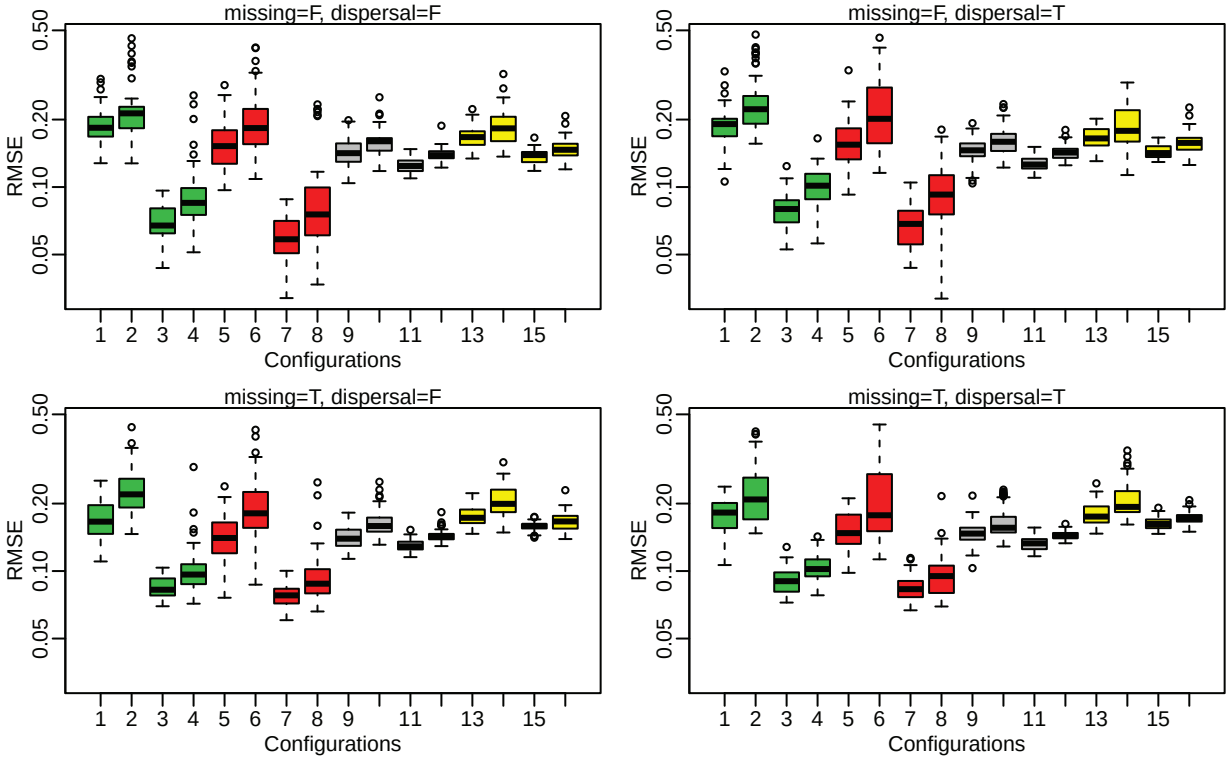


Figure S5: Boxplots of RMSE for simulated species 6. Same caption as Figure S1.

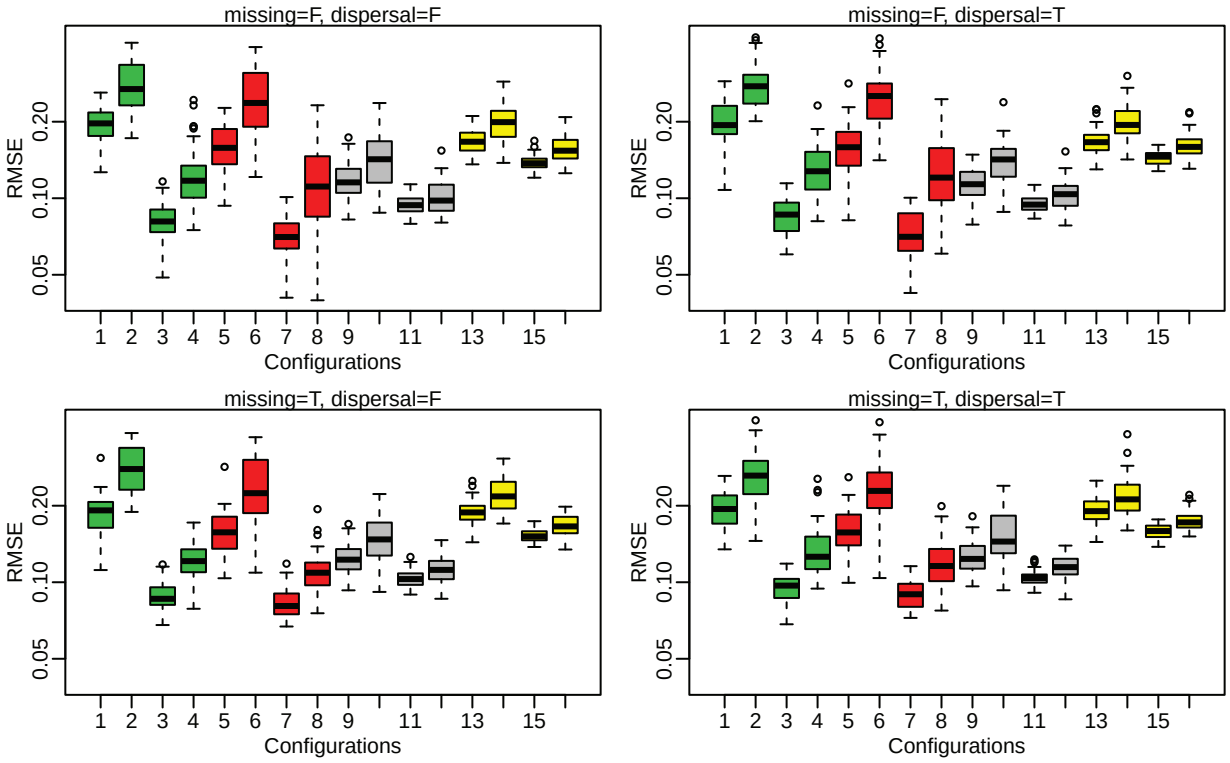


Figure S6: Boxplots of RMSE for simulated species 7. Same caption as Figure S1.

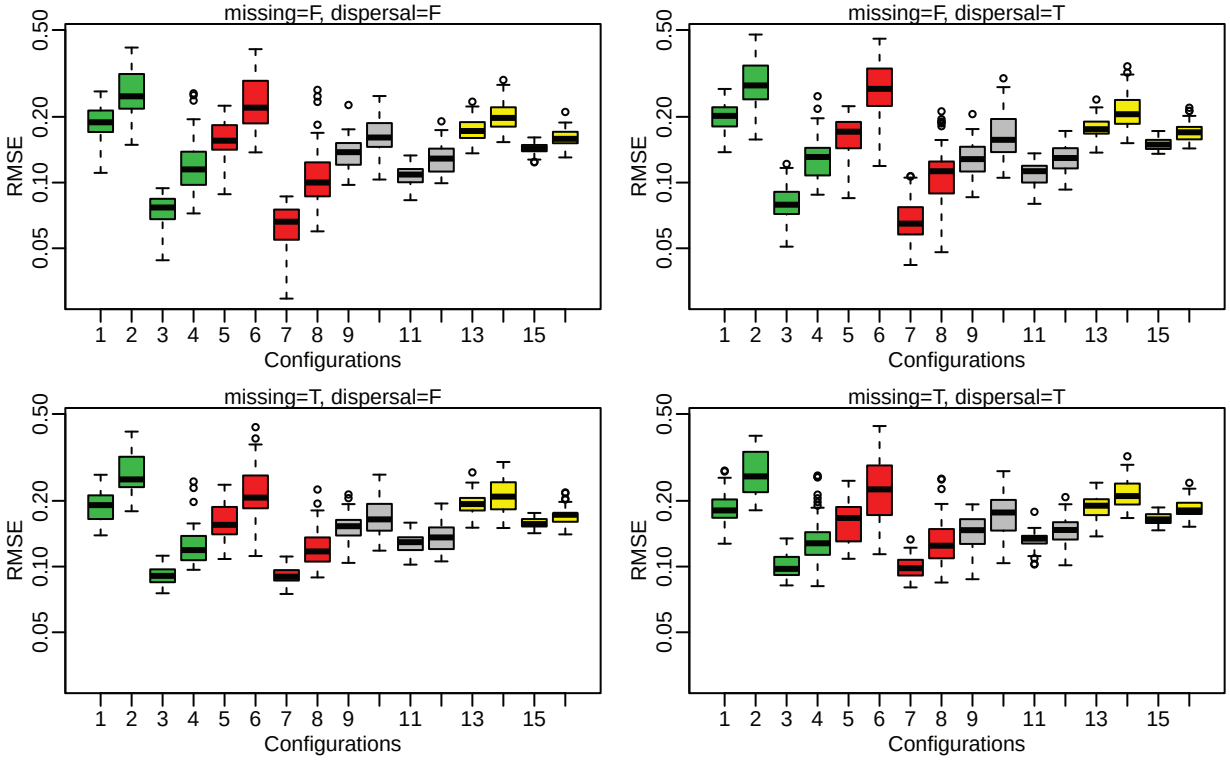


Figure S7: Boxplots of RMSE for simulated species 8. Same caption as Figure S1.

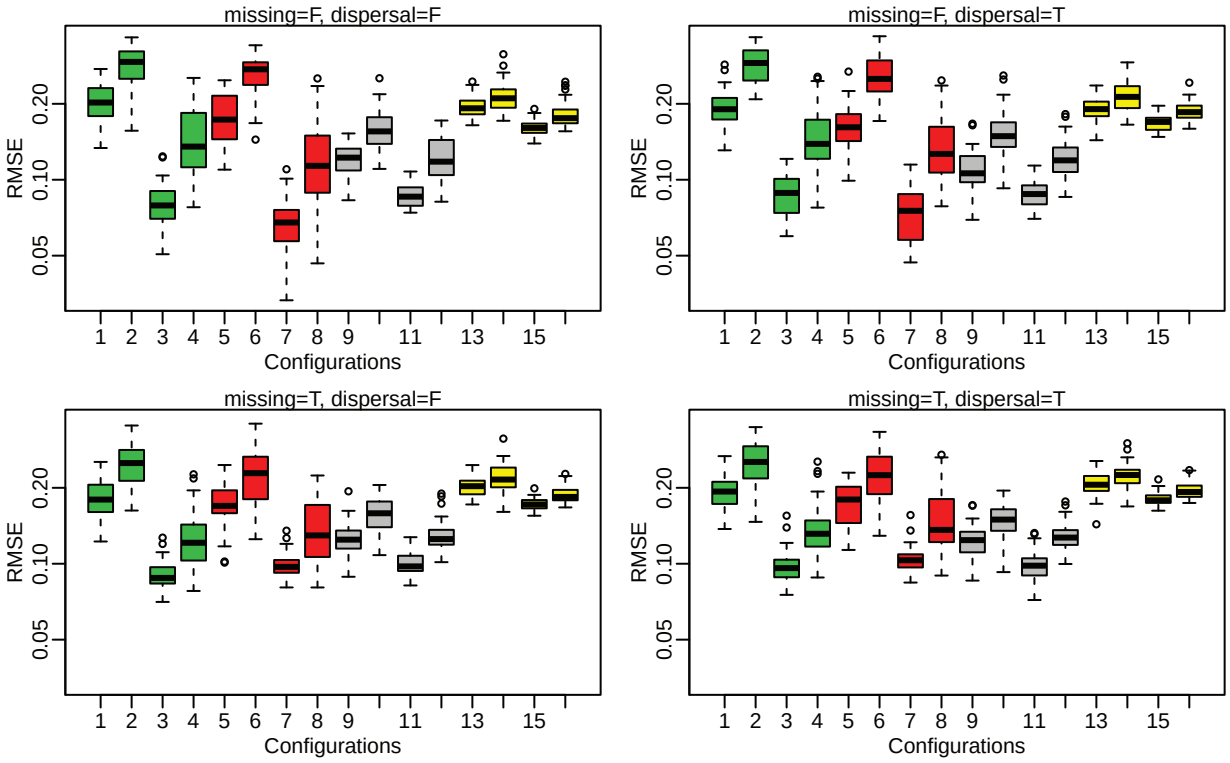


Figure S8: Boxplots of RMSE for simulated species 9. Same caption as Figure S1.

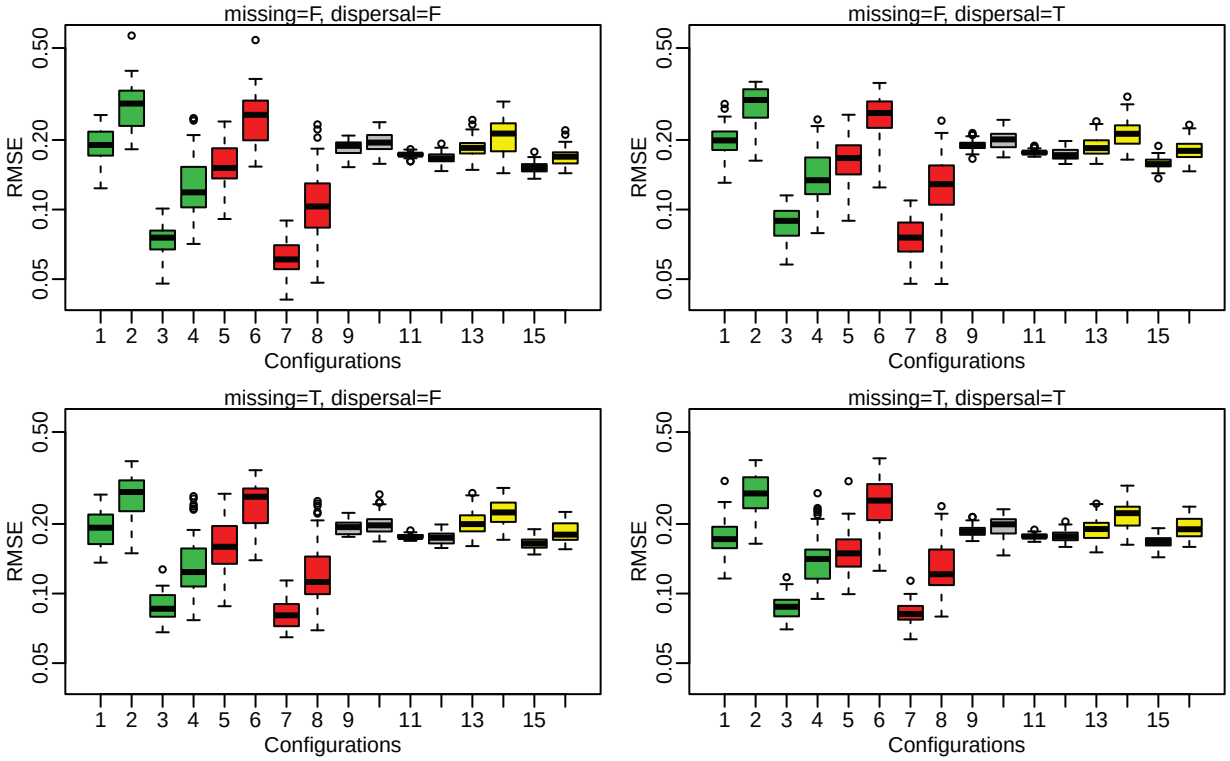


Figure S9: Boxplots of RMSE for simulated species 10. Same caption as Figure S1.

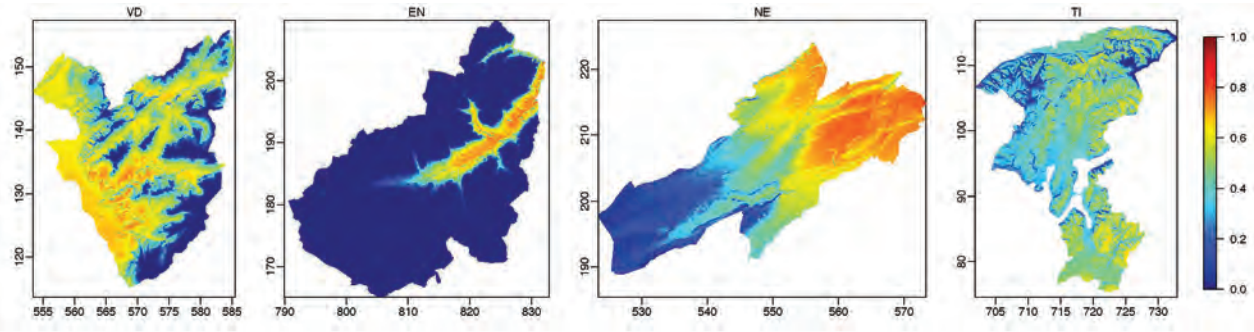


Figure S10: Theoretical probabilities of presence for simulated species 1.

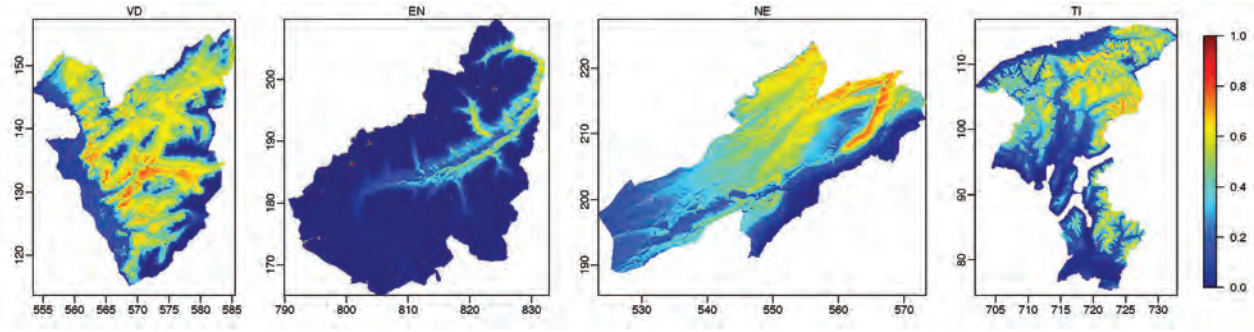


Figure S11: Theoretical probabilities of presence for simulated species 2.

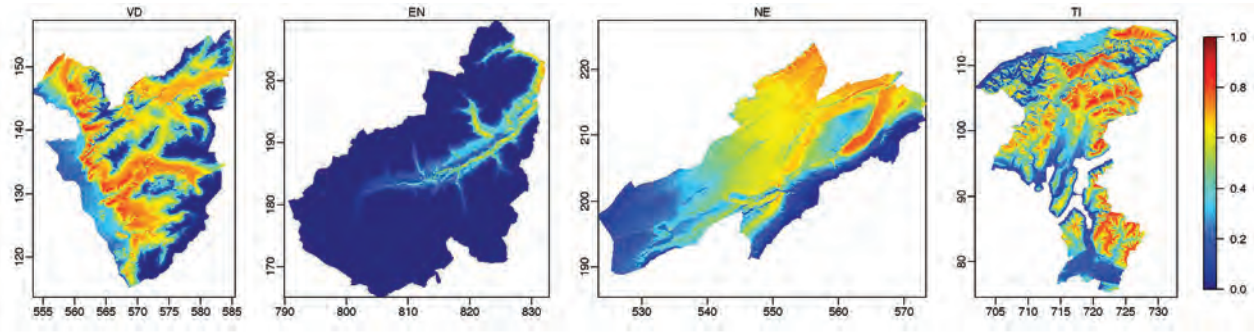


Figure S12: Theoretical probabilities of presence for simulated species 3.

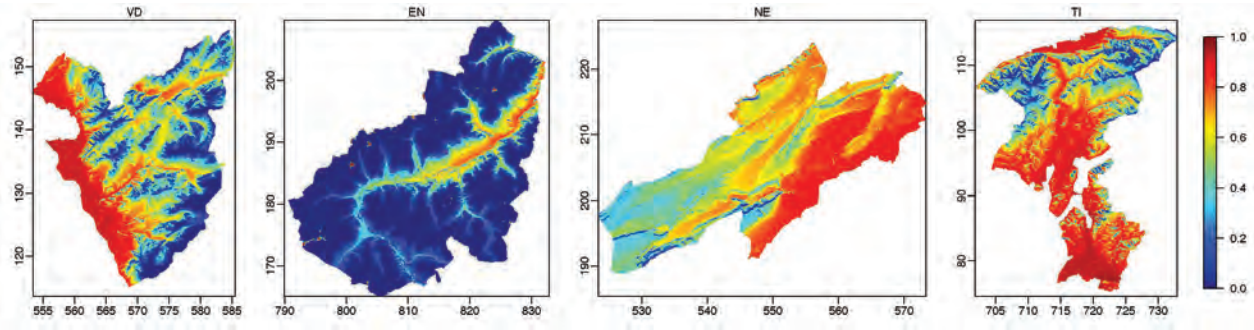


Figure S13: Theoretical probabilities of presence for simulated species 4.

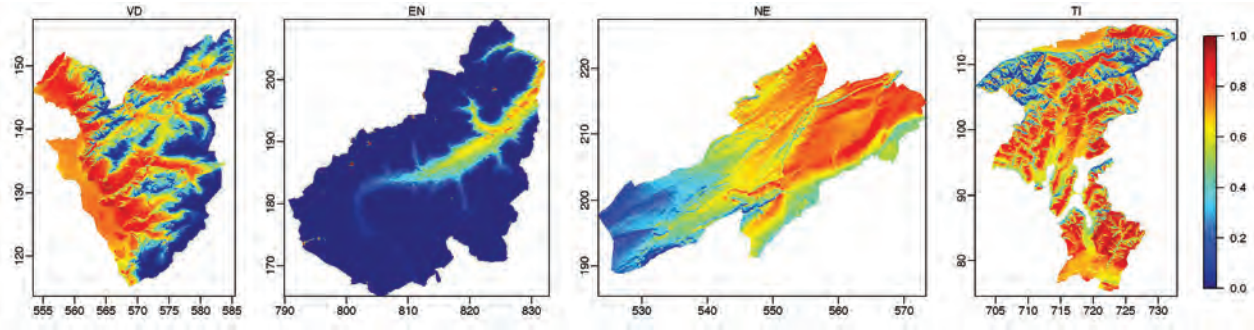


Figure S14: Theoretical probabilities of presence for simulated species 5.

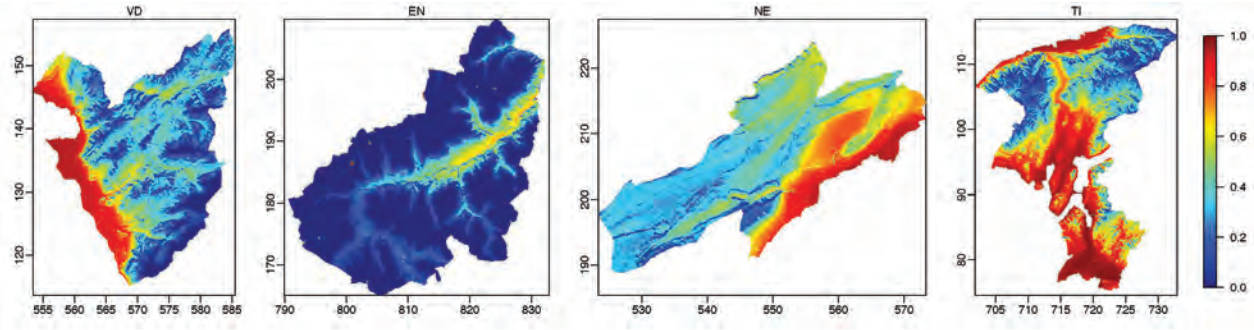


Figure S15: Theoretical probabilities of presence for simulated species 6.

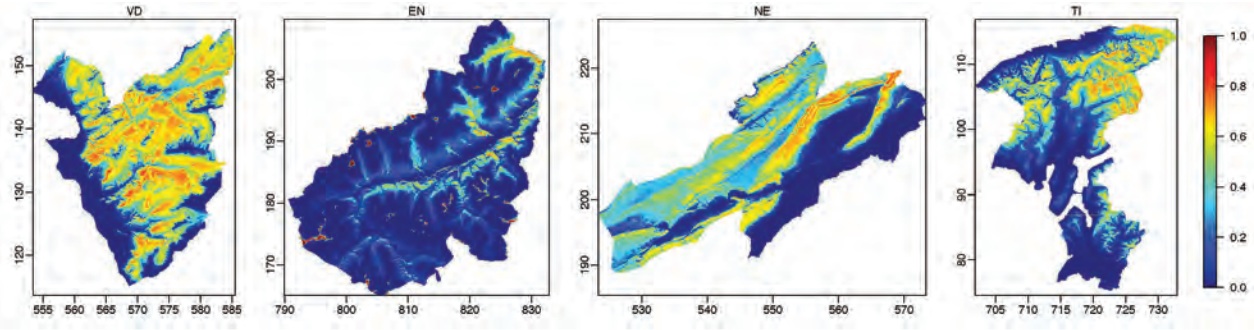


Figure S16: Theoretical probabilities of presence for simulated species 7.

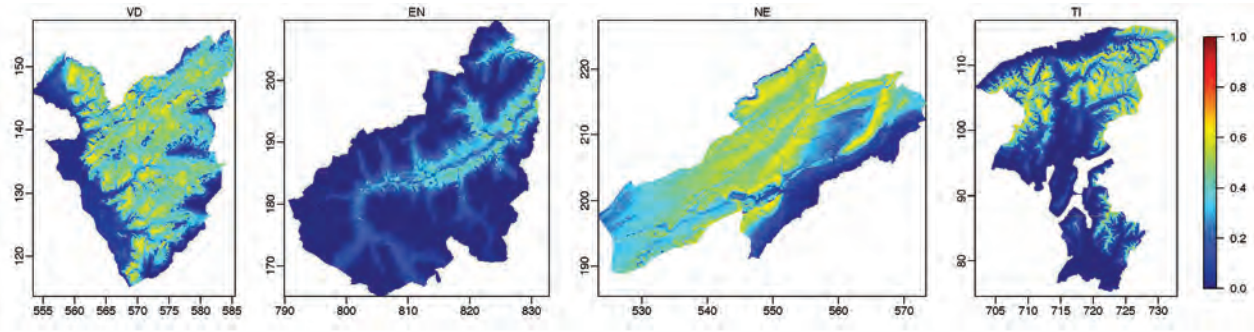


Figure S17: Theoretical probabilities of presence for simulated species 8.

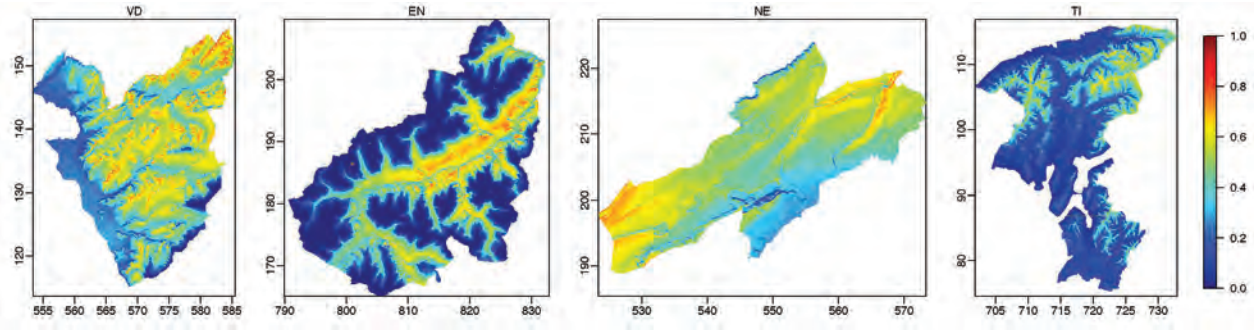


Figure S18: Theoretical probabilities of presence for simulated species 9.

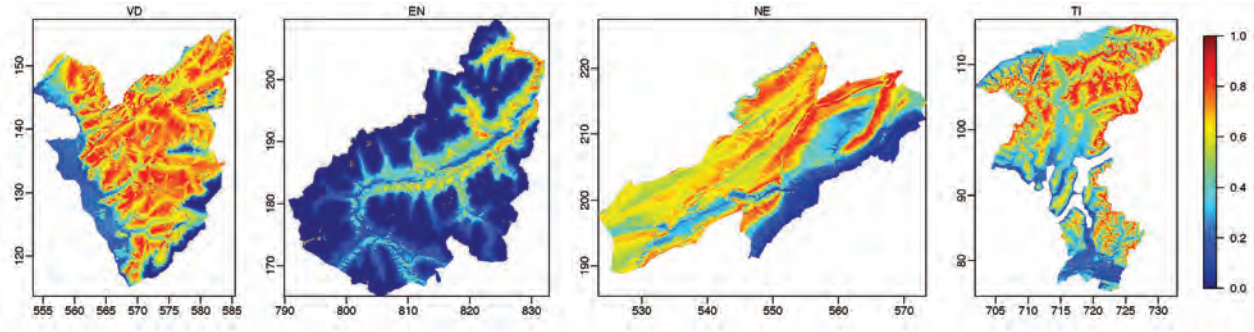


Figure S19: Theoretical probabilities of presence for simulated species 10.

References

Albert, C.H., Yoccoz, N.G., Edwards, T.C., Graham, C.H., Zimmermann, N.E. & Thuiller, W. (2010) Sampling in ecology and evolution – bridging the gap between theory and practice. *Ecography*, **33**, 1028–1037.

Austin, M.P. & Adomeit, E.M. (1991) Sampling strategies costed by simulation. C.R. Margules & M.P. Austin, eds., *Nature Conservation: cost effective biological surveys and data analysis*, pp. 167–175. Commonwealth Scientific & Industrial Research (CSIRO), Dickson, Australia.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481–497.

Cablk, M., D., W. & R., K.A. (2002) Assessment of spatial autocorrelation in empirical models in ecology. J. Scott, P. Heglund, M. Morrison, J. Haufler, M. Raphael, W. Wall & F. Samson, eds., *Predicting Species Occurrences: Issues of Accuracy and Scale*, pp. 429–440. Island Press, Covelo, CA.

Caruana, R. & Niculescu-Mizil, A. (2004) Data mining in metric space: an empirical analysis of supervised learning performance criteria. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '04, pp. 69–78. ACM, New York, NY, USA.

Dormann, C.F. (2007) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, **16**, 129–138.

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

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

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M. & E. Zimmermann, N. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

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.

Freeman, E.A. & Moisen, G.G. (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, **217**, 48–58.

Graham, C.H., Elith, J., Hijmans, R.J., Guisan, A., Townsend Peterson, A. & Loiselle, B.A. (2008) The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, **45**, 239–247.

Graham, M.H. (2003) Confronting Multicollinearity in Ecological Multiple Regression. *Ecology*, **84**, 2809–2815.

Guisan, A., Zimmermann, N.E., Elith, J., Graham, C.H., Phillips, S. & Peterson, A.T. (2007a) What Matters for Predicting the Occurrences of Trees: Techniques, Data, or Species' Characteristics? *Ecological Monographs*, **77**, 615–630.

Guisan, A., Graham, C.H., Elith, J. & Huettmann, F. (2007b) Sensitivity of predictive species distribution models to change in grain size. *Diversity and Distributions*, **13**, 332–340.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773–785.

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., Helfer, V. & Metral, F. (2001) Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, **145**, 111–121.

Hortal, J., Borges, P.A.V. & Gaspar, C. (2006) Evaluating the performance of species richness estimators: sensitivity to sample grain size. *Journal of Animal Ecology*, **75**, 274–287.

Jiménez-Valverde, A., Lobo, J.M. & Hortal, J. (2009) The effect of prevalence and its interaction with sample size on the reliability of species distribution models. *Community Ecology*, **10**, 196–205.

Johnson, C.J. & Gillingham, M.P. (2008) Sensitivity of species-distribution models to error, bias, and model design: An application to resource selection functions for woodland caribou. *Ecological Modelling*, **213**, 143–155.

Kadmon, R., Farber, O. & Danin, A. (2003) A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, **13**, 853–867.

Kadmon, R., Farber, O. & Danin, A. (2004) Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecological Applications*, **14**, 401–413.

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

Liu, C., White, M. & Newell, G. (2011) Measuring and comparing the accuracy of species distribution models with presence–absence data. *Ecography*, **34**, 232–243.

M McPherson, J. & Jetz, W. (2007) Effects of species' ecology on the accuracy of distribution models. *Ecography*, **30**, 135–151.

Mason, S.J. & Graham, N.E. (2002) Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society*, **128**, 2145–2166.

McPherson, J., Jetz, W. & Rogers, D. (2006) Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions–possibilities and limitations. *Ecological Modelling*, **192**, 499–522.

McPherson, J.M., Jetz, W. & Rogers, D.J. (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.

Meynard, C.N. & Quinn, J.F. (2007) Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *Journal of Biogeography*, **34**, 1455–1469.

Moisen, G. & Frescino, T. (2002) Comparing five modelling techniques for predicting forest characteristics. *Ecological Modelling*, **157**, 209–225.

Naimi, B., Skidmore, A.K., Groen, T.A. & Hamm, N.A.S. (2011) Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. *Journal of Biogeography*, **38**, 1497–1509.

Nakagawa, S. & Schielzeth, H. (2013) A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, **4**, 133–142.

Peterson, A.T. & Cohoon, K.P. (1999) Sensitivity of distributional prediction algorithms to geographic data completeness. *Ecological Modelling*, **117**, 159–164.

Phillips, S.J., Dudík, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Ferrier, S. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181–197.

Pottier, J., Dubuis, A., Pellissier, L., Maiorano, L., Rossier, L., Randin, C.F., Vittoz, P. & Guisan, A. (2013) The accuracy of plant assemblage prediction from species distribution models varies along environmental gradients. *Global Ecology and Biogeography*, **22**, 52–63.

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

Reese, G.C., Wilson, K.R., Hoeting, J.A. & Flather, C.H. (2005) Factors affecting species distribution predictions: a simulation modeling experiment. *Ecological Applications*, **15**, 554–564.

Segurado, P., Araújo, M.B. & Kunin, W.E. (2006) Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, **43**, 433–444.

Segurado, P. & Araújo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.

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

Tate, R.F. (1954) Correlation Between a Discrete and a Continuous Variable. Point-Biserial Correlation. *The Annals of Mathematical Statistics*, **25**, 603–607.

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

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D. & Kadmon, R. (2007) A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions*, **13**, 397–405.

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

Varin, C., Reid, N. & Firth, D. (2011) An overview of composite likelihood methods. *Statistica Sinica*, **21**, 5–42.

Vittoz, P. & Engler, R. (2007) Seed dispersal distances: a typology based on dispersal modes and plant traits. *Botanica Helvetica*, **117**, 109–124.

Wisz, M.S. & Guisan, A. (2009) Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, **9**, 8.

Wisz, M.S., Hijmans, R.J., Peterson, A.T., Graham, C.H., Guisan, A. & NCEAS Predicting Species Distributions Working Group (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

Zaniewski, A., Lehmann, A. & Overton, J. (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.

Global Ecology and Biogeography
A Journal of Macroecology

Residence time, expansion toward the equator in the invaded range and native range size matter to climatic niche shifts in non-native species

Journal:	*Global Ecology and Biogeography*
Manuscript ID:	GEB-2013-0141.R3
Manuscript Type:	Research Papers
Date Submitted by the Author:	n/a
Complete List of Authors:	Li, Yiming; Institute of Zoology, The Chinese Academy of Sciences, Key Laboratory of Animal Ecology and Conservation Biology Liu, Xuan; Institute of Zoology, Chinese Academy of Sciences, Key Laboratory of Animal Ecology and Conservation Biology Li, Xianping; Institute of Zoology, The Chinese Academy of Sciences, Key Laboratory of Animal Ecology and Conservation Biology Petitpierre, Blaise; University of Lausanne, Department of Ecology and Evolution Guisan, Antoine; University of Lausanne, Department of Ecology and Evolution
Keywords:	ecological niche models, non-native species, realized climatic niche, niche shifts, expansion towards equator, native range size, residence time, climate match

SCHOLARONE™
Manuscripts

Residence time, expansion toward the equator in the invaded range and native range size matter to climatic niche shifts in non-native species

Yiming Li[1,a *], Xuan Liu[1,a], Xianping Li [1,2,a], Blaise Petitpierre[3] and Antoine Guisan[3]

[1]*Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China.*

[2]*University of Chinese Academy of Sciences, 19 Yuquan Road, Shijingshan, Beijing, 100049, China*

[3]*Department of Ecology and Evolution, University of Lausanne, Biophore Building, CH-1015 Lausanne, Switzerland*

[a]co-first authors contributed equally to this work.

*Corresponding author: YL, liym@ioz.ac.cn

XLIU: liuxuan@ioz.ac.cn, XL: lixianping@ioz.ac.cn, BP: bpetitpierre@gmail.com, AG: antoine.guisan@unil.ch

Running title: Climatic niche shifts and their determinants

Keywords: ecological niche models (ENMs), non-native species, realized climatic niche, niche shifts, expansion towards equator, native range size, residence time, climate match

The number of words in the Abstract: 268

The number of words in main body of the paper: 5095 words, with two tables and three figures.

The number of references: 50

For Peer Review

Abstract

Aim Identifying climatic niche shifts and their drivers is important to accurately predict the risk of biological invasions. The niches of non-native plants and birds have recently been assessed in large-scale multi-species studies, but such large-scale tests are lacking for non-native reptiles and amphibians (herpetofauna). Furthermore, little is known about the factors contributing to niche shifts when they occur. Based on the occurrence of 71 reptile and amphibian species, we compared native and non-native realized niches in 101 invaded ranges at a worldwide scale and identified the factors that affect niche shifts.

Location The world except the Antarctic.

Methods We assessed climatic niche dynamics in a gridded environmental space allowing the quantification of niche overlap and expansion into climatic conditions not colonized by the species in their native range. We analyzed the factors affecting niche shifts using a model averaging approach based on generalized linear mixed-effects models.

Results Approximately 57% of the invaded ranges (51% for amphibians and 61% for reptiles) showed niche shifts (≥10% expansion in the realized climatic niche). Island endemics, species introduced to Oceania and invaded ranges outside the native biogeographic realm showed a higher proportion of niche shifts. Niche shifts were more likely for species that had smaller native range sizes, were introduced earlier into a new range or invaded areas located at lower latitudes than the native range.

Main conclusions The proportion of niche shifts for non-native herpetofauna was

For Peer Review

higher than those for Holarctic non-native plants and European non-native birds. The 'climate matching hypothesis' should be used with caution for species shifting their niche because it could underestimate the risk of their establishment.

For Peer Review

INTRODUCTION

Niche conservatism refers to the tendency of a species' ecological niche to be conserved over space and time, an assumption that is increasingly considered in evolutionary, ecological and conservation studies (Peterson, 2011). In particular, it is a pivotal assumption for ecological niche models (ENMs), which depict Grinnellian species niches by correlating species' geographic occurrences with environmental variables (i.e., 'climate-matching' with climate variables only) and use them to project the potential distribution of species in time and space (Guisan & Thuiller, 2005; Pearman *et al.*, 2008; Soberón & Nakamura, 2009; Peterson *et al.*, 2011; Araújo & Peterson 2012; Liu *et al.*, 2013). Nevertheless, there is still considerable debate on the climate niche conservatism of species (Losos, 2008; Wiens *et al.* 2010; Peterson, 2011; Peterson and Nakazawa 2008; Pearman *et al.* 2013).

Non-native species offer excellent model systems for examining niche conservatism and evolution within a short time scale through comparisons of climate attributes between the native and invaded ranges (Sax *et al.* 2007; Petterson, 2011; Schulte *et al.* 2012). Several studies have identified realized climatic niche shifts for non-native species (Fitzpatrick *et al.*, 2007; Broennimann & Guisan, 2008; Beaumont *et al.*, 2009; Gallagher *et al.*, 2010), demonstrating the limited capacity of 'climate matching' approaches to predict the potential geographic extent of invasion. In this context, a robust framework serving to formalize, quantify and statistically test for niche shifts of non-native species in environmental space has recently been developed (Warren et al. 2008; Broennimann *et al.*, 2012; Petitpierre *et al.*, 2012).

Assuming that a species has colonized all suitable conditions in its native range, the

For Peer Review

realized niche quantified in its invaded range can be divided into three parts (Petitpierre *et al.*, 2012): niche stability, the part of the niche in which the species occurs in both its native and invaded ranges; niche expansion, where the species newly occurs in the invaded range; and niche unfilling, where the species occurs only in its native range. As biological invasions are recent and ongoing processes, niche unfilling is likely due to dispersal limitations in the invaded range. Thus, only niche expansion toward a climate that is available but not colonized in the native range (i.e., an analog climate) can be unambiguously considered as the niche shifts. Based on such a framework, two studies have suggested that climatic niche shifts are rare for non-native plants (<15% of 50 species displaying niche shifts, i.e., ≥10% expansion in realized climatic niche; Petitpierre *et al.*, 2012) and birds (29% of 28 species showing niche shifts; Strubbe *et al.*, 2013). Whether the same patterns of niche shifts are found in other organisms remains unknown.

Identifying the factors that contribute to niche shifts between invaded and native ranges will not only help to clarify the debate but will also be of fundamental importance for our understanding of species distributions and their responses to changing environments, e.g., through ENM predictions. It has been hypothesized that realized niche shifts can result from both 1) changes in dispersal limitations and biotic interactions between native and invaded ranges and 2) introduction history and rapid evolution (Pearman *et al.*, 2008; Alexander & Edwards, 2010). First, for species with restricted native ranges (such as island and mountain endemics), realized niche shifts most likely result from dispersal limitations (including extrinsic dispersal limitation,

For Peer Review

e.g., barriers to dispersal, and intrinsic dispersal limitation, e.g., a species' ability to disperse) in the native range (Alexander & Edwards, 2010). Realized niche shifts would similarly be more likely for species introduced into lower-latitude areas (Alexander & Edwards, 2010). It is hypothesized that the low-latitude boundary of a species' geographical range is determined by its tolerance to biotic pressures from more complex biotic interactions (Darwin, 1859; Normand *et al.*, 2009). Because species richness tends to be greater toward the equator, greater competition and predation pressures would be expected in low-latitude areas, resulting in more complex biotic interactions there. A low-latitude boundary expansion in the invaded range relative to that in the native range may thus contain climates that were not colonized in the native range due to limiting biotic pressures, which are more likely to cause realized niche shifts.

Second, it is hypothesized that the realized climatic niche would tend to shift as a function of the residence time in the invaded range (Gallagher *et al.*, 2010; Peterson, 2011). The invaded range will expand with increasing residence time (Wilson *et al.*, 2007; Williamson *et al.*, 2009), thus becoming increasingly likely to include those climatic habitats that were excluded by biotic interactions and dispersal limitations in the native range. With increasing residence time, a species would also be more likely to evolve adaptations to novel climate conditions not found in the native range (non-analog) and to expand its fundamental niche (Peterson, 2011). In addition, multiple introductions of a non-native species from different source populations in the native range may facilitate the admixture of previously isolated native populations

For Peer Review

and increase the genetic variation in invading populations (Kolbe *et al.*, 2004). Such added genetic variation may promote the capacity of a population in the invaded range to respond to selection in new environments, favoring the occurrence of fundamental niche shifts (Pearman *et al.*, 2008; Alexander & Edwards, 2010).

Although ecological and evolutionary drivers of niche changes cannot easily be disentangled from observational data, a first necessary step using such data is, nevertheless, to assess whether niche changes occur and in what proportions. However, niche change studies across many species and over large areas remain scarce. Here, we assessed realized climatic niche shifts for non-native terrestrial reptile and amphibian species (non-native herpetofauna) by comparing the realized niches between invaded and native ranges, using empirical distribution data and identifying factors that could affect niche shifts in invaded ranges.

As ectothermic organisms, reptiles and amphibians depend on external heat to increase their body temperature and become active. Climate affects all important aspects of reptile and amphibian biology, including growth, development, foraging and the timing of hibernation and breeding (Vitt & Caldwell, 2009). Both temperature and precipitation are considered to have marked effects on the distribution, range size and species richness patterns of reptiles and amphibians (Araújo *et al.*, 2008; Aragón *et al.*, 2010; Whitton *et al.*, 2012). Many reptile and amphibian species have been introduced into new ranges via the pet and food trade, deliberate introduction for personal aesthetic pleasure or aquaculture and the deliberate or accidental release of pets and aquarium specimens (Kraus, 2009). However, unlike the introduction of

For Peer Review

non-native plants and mammals, the widespread introduction of herpetofauna is only recent, and the introduction history of most species has been documented in relatively great detail (Kraus, 2009). These provide a unique opportunity to examine the effects of the introduction history (the number of introduction events and the residence time since arrival) on niche shifts. In this paper, we ask the following questions:

1) How frequent and important are realized climatic niche shifts within non-native herpetofauna in invaded ranges at a global scale?

2) Which factors promote realized climatic niche shifts in the invaded ranges? We tested for possible associations between niche shifts and native range size, island endemic identity, the direction of latitudinal expansion in the invaded range, the number of introduction events and the residence time since arrival.

Assessing and explaining the magnitude of realized climatic niche shifts would allow a better understanding of species responses towards changing environments and would facilitate more efficient management strategies for invasive species.

METHODS

Data on occurrences of non-native herpetofauna

We obtained data on successfully introduced amphibian and reptile species from the database of Kraus (2009), which is widely used in studies of non-native herpetofauna (Tingley *et al.*, 2011; Van Wilgen & Richardson, 2012). Subsequently, we validated the data and determined the final species list (Text S1). If taxonomic inconsistencies were found, they were resolved by further literature review.

For Peer Review

Occurrence data for each species were obtained from various databases and published references (Table S1). Here, native and invaded ranges within the same biogeographic realm can be clearly distinguished (Kraus, 2009). We therefore distinguished, in our dataset, invaded ranges within the same biogeographic realm from those outside the native biogeographic realm. A total of 68,459 location records in native ranges (range: 14-8741 records) and 15,973 records in introduced ranges (range: 10-2740 records) were collected for 71 species of non-native reptiles and amphibians. Previous studies have been conducted at a coarse resolution (e.g., grid cell of 30 arc-minutes ≈ 50 × 50 km; Petitpierre *et al.*, 2012; Strubbe *et al.*, 2013). Given that certain fine-scale niche shifts might not be detected at such a resolution (Petitpierre *et al.*, 2012), we integrated our data into grid cells at both a finer resolution (10 arc-minutes ≈ 16 × 16 km) and the same coarse resolution (30 arc-minutes) after removing duplicate observations in the same geographic cell.

Climatic variables

Eight climatic variables from the WorldClim database, widely used in large-scale studies of species distributions (Hijmans *et al.*, 2005), were used to depict the realized climatic niche of species: mean diurnal range of temperature, annual range of temperature, minimum temperature of the coldest month, mean temperature of the warmest quarter, annual precipitation, precipitation seasonality, precipitation in the driest quarter and precipitation in the warmest quarter. These variables describe the energy and water factors that define the primary physical requirements of amphibians and reptiles and are widely used for determining the geographical distributions of

For Peer Review

amphibians and reptiles (Araújo *et al.*, 2008).

Quantifying the dynamics of realized climatic niches

The extent of the study area has important effects on niche comparisons. The spatial and temporal extent of the study area should be sufficiently broad to allow the optimal coverage of the spatial footprint of all suitable environmental conditions and to incorporate the range of factors that typically affect species ranges, such as geographical barriers, climates, historical events and phylogeographical phenomena (Barve *et al.*, 2011; Soberón & Peterson, 2011). We used biogeographic realm(s) to define native and invaded extent(s) following Olson *et al.* (2001)'s definition: Afrotropics (including Madagascar), Australasia, Indo-Malay, Nearctic, Neotropics, Palearctic and Oceania. Several species were distributed in two or more native realms. The native realm for such a species was defined as the realm where the midpoint of its latitudinal and longitudinal ranges was located.

We quantified the dynamics of climatic niches in analog climates (i.e., the climates that are available in both native and invaded extents; see Fitzpatrick & Hargrove, 2009) between the native and the invaded ranges for every species at both coarse and finer resolutions. Non-analog climates were excluded from the analyses because niche extrapolations in this case are difficult to interpret (Fitzpatrick & Hargrove, 2009; Guisan *et al.*, 2012; Petitpierre *et al.*, 2012). We gridded the species occurrences in the native and invaded ranges and, to retain sufficient statistical power, we only included in our analyses species occurring in ≥ 10 grid cells in either their

For Peer Review

native (range: 11 - 1200 grid cells) or invaded (range: 10 - 695 grid cells) ranges.

Species data were projected onto the first two axes of a principal component analysis (PCA), depicting a multivariate climatic space calculated with the eight climatic variables (Fig. S1). Following previous studies (Broennimann *et al.*, 2012; Petitpierre *et al.*, 2012; Strubbe *et al.*, 2013), we did not include additional axes because the first two already explained a large proportion of the total climate variation (Fig. S1, Tables S2-3). We used two PCA calibration approaches (both referred to as PCA_{env} in Broennimann *et al.*, 2012). A first PCA was calibrated on climate factors distributed in the native extent only, and a second was calibrated on both native and invaded extents. Correlations between the first two PCA axes and climate factors vary depending on the extent used for the PCA calibration (Table S3). Species occurrences were then transformed into species density using a kernel smoother in the gridded PCA environmental space (resolution = 100*100) (Broennimann *et al.*, 2012). This approach reduces the risk that a difference between the number of native and introduced records would cause an analytical bias in our results. The same gridding and smoothing procedure was applied to the entire climate available in the native and invaded extents. This approach allowed the definition of species occupancy by correcting species densities through the incorporation of the differences in environmental availability between native and invaded ranges (Broennimann *et al.*, 2012). Niche overlap between native and invaded ranges was quantified through Schoener's D (Warren *et al.*, 2008; Broennimann *et al.*, 2012), varying between 0 (no overlap) and 1 (total overlap). This approach has been suggested as the most accurate

For Peer Review

to evaluate realized niche dynamics among numerous methods (Broennimann *et al.*, 2012). Additionally, it allows testing for niche conservatism through a one-sided niche similarity test based on D (Broennimann *et al.*, 2012).

D does not reveal the detailed nature of the niche changes. Therefore, we also quantified three more specific components of niche change: unfilling, niche stability and niche expansion (Petitpierre *et al.*, 2012). Unfilling corresponds to the portion of niche space colonized only in the native range, stability to the portion colonized in both native and invaded ranges and expansion to the portion colonized only in the invaded range. These indices were measured in the climatic space shared between the native and invaded extents to avoid detection of niche shifts due to climatic non-availability in the native range. Note that in both the native and invasive extents, marginal climates with densities smaller than 25% were not included to reduce the heterogeneity in climate availability between native and invaded extents. The results were similar for different proportions of the intersection of the species densities (75, 80, 85, 90, 95 and 100%; Fig. S2). We used the same R functions as Petitpierre *et al.* (2012) for the whole procedure.

Factors influencing realized climatic niche shifts

Range size was measured as the number of 10 arc-minute occupied grid cells, using Hawth's Tools in ArcGIS (Beyer, 2010; Gallagher *et al.*, 2010). Island endemics, i.e., species whose native range only included islands, were represented by a binary variable. The number of successful introduction events and the year of the first

For Peer Review

successful introduction of non-native species were obtained from Kraus (2009). Residence time was determined as the number of years since the first successful introduction in the invaded biogeographic realms. Expansion toward the equator or toward the pole (hereafter, equatorward and poleward expansion, respectively) in the invaded realms was defined by quantifying cases where the high and low latitude distribution limits in the invaded extents were, respectively, higher and lower than in the native extent. In all other cases, the expansion toward the equator and toward the poles was set to zero (no expansion).

Statistical Analyses

We used multi-model inference based on information theory (Burnham & Anderson, 2002) to investigate the effects of these factors on the occurrence of niche shifts (≥10% expansion = 1, others = 0). This approach allows more reliable inferences from an entire set of models and is preferable to selecting a single best model (Burnham & Anderson, 2002). The full model is a generalized linear mixed-effects model (GLMM) with a logit link and binomial error distribution, with niche shifts as the response variable and the six factors as predictor variables. We included the invaded and native realms as a random factor (Gallagher *et al.*, 2010); all other variables were treated as fixed effects. To avoid potential statistical issues arising from phylogenetic non-independence, we evaluated the phylogenetic signal in the niche shifts for amphibians and reptiles, respectively (Table S4). We obtained data on phylogenetic distances from published supertrees (Table S4). We estimated Pagel's

For Peer Review

λ phylogenetic signal using a maximum likelihood approach, and tested for its significance compared to a Pagel's λ of zero with the likelihood ratio test. We performed these analyses using the package phytools (see Table S4). As we did not find a significant relationship between phylogeny and climatic niche shifts (Table S4), we did not consider phylogeny any further in models.

Because we were interested in the relative importance of an individual factor, we excluded models with interaction terms from the calculation of relative variable importance and thus created 63 models (2^6-1) representing all possible combinations of six predictor variables. We compared the alternative models using the Akaike information criterion adjusted for small samples (AIC_c). We calculated the relative importance of a predictor by summing the Akaike weights across all of the models in which the predictor appeared. We calculated the model-averaged parameter estimate for each predictor and its variances with Akaike weights (Table S5) (Burnham & Anderson, 2002). We reported the best models that were within 2 AIC_c units as the highest ranked models (i.e., $\Delta AIC_c \leq 2$; Burnham & Anderson, 2002). We used the lmer function in the lme4 package to perform the GLMMs analysis and used the dredge and model.avg functions in the MuMIn package to perform the model averaging analysis. All analyses were conducted in R (R Development Core Team, 2011). The full analytical procedure is shown in Fig. S3. The data used are shown in Table S6.

RESULTS

For Peer Review

Realized climatic niche shifts

Our sample at the fine resolution contained 101 invaded ranges for 71 species of non-native reptiles (46 species) and amphibians (25 species), including 53 invaded ranges within the native realm and 48 invaded ranges outside the native realm (Table S2). At the coarse resolution, the sample size was reduced to 58 invaded ranges for 36 species.

Globally, there were no differences in the niche shifts, niche overlap and niche unfilling in the invaded ranges at the same resolution between different realms or between different resolutions based on the same realm (Tables S7A-C). These results indicated that the realized climatic niche dynamics between invaded and native ranges for non-native herpetofauna were highly consistent despite the use of methods based on different realms or different resolutions. Therefore, we only report here the results of niche shifts quantified by PCA calibrated on climate availability within the native range realms at the fine resolution due to the larger sample size of invaded ranges and the consistent use of realm (the results of niche shifts by PCA calibrated on native and invaded ranges were similar and are shown in Tables S2, S6-8).

Realized climatic niche conservatism (i.e. niche stability) between the invaded and native ranges was observed for 50% (50/101) of these invaded ranges at the fine resolution ($P \leq 0.05$ for the similarity test) (Table S2). Niche unfilling (≥10%) was observed for 80% (80/101) of the invaded ranges. Fig. 1 illustrates the climate niche dynamics between the invaded and native ranges for two notorious globally invasive species.

For Peer Review

At the global scale, the proportion of climatic niche shifts in all invaded ranges was high. Out of the 101 invaded ranges (Table S2; Fig. 2a; Fig. S4), 58 showed greater than 10% niche expansion, with 18 ranges displaying niche shifts (≥10% expansion) for amphibians and 40 for reptiles. *Discoglossus pictus,* introduced from the Palearctic into the same province outside its native range, showed 92.8% expansion, the highest expansion found in an invaded range for amphibians. *Caiman crocodilus*, introduced from the Neotropics into the Nearctic, displayed 90.5% expansion, the highest found for reptiles (Table S2). The proportion of niche shifts in invaded ranges at the global scale did not differ between reptiles and amphibians (Chi-square test, χ^2 = 0.780, df = 1, P = 0.377). No difference between amphibians and reptiles in the proportion of species showing niche shifts in one or more realms was detected at the species level (χ^2 = 0.132, P = 0.716). However, invaded ranges within the native realm showed a lower proportion of niche shifts than those in different realms (χ^2 = 4.239, P = 0.04) (Fig. 2b). Furthermore, island endemics showed a higher proportion of niche shifts than other species (χ^2 = 6.964, P = 0.008) (Fig. 2b).

At the biogeographic realm scale, the proportion of niche shifts in invaded ranges varied strongly, ranging from 0 to 75% for amphibians and from 0 to 100% for reptiles (Fig. 3). No difference was found in the proportion of niche shifts among the seven realms (Kruskal-Wallis test, χ^2 = 10.986, df = 6, P = 0.089). However, the invaded ranges for species introduced to Oceania showed more niche shifts than those for species introduced to the combination of the other six realms (Chi-square test, χ^2 = 6.777, df = 1, P = 0.009).

For Peer Review

Factors affecting climatic niche shifts

Controlling for the realm into which a species was introduced and the realm from which the species was native, the six best models (i.e., $\Delta AIC_c \leq 2$) contained all predictors (Table 1). Native range size, residence time and equatorward expansion in the invaded range entered into each of these models. The proportion of the variation (pseudo-R^2) explained by predictors in models ranged from 0.64 to 0.67.

The model averaging analysis showed that native range size (1.0), residence time (1.0) and equatorward expansion (1.0) had high relative importance values for the realized climatic niche shifts in the invaded ranges (Table 2). The other three variables (poleward expansion, island endemic identity and number of introduction events) showed lower relative importance values ranging from 0.38 to 0.53. The amplitude of niche shifts increased (positive parameter estimates) with increasing residence time, equatorward expansion, poleward expansion, island endemic identity and number of introduction events but decreased (negative parameter estimates) with increasing native range size.

The results were similar when the PCA was calibrated using extent data comprising both native and invaded ranges (Table S8A) or when the earliest introduction (occurred 2212 years ago for *Macroprotodon cucullatus*, Table S6) was excluded from the sample (Tables S8B-C).

DISCUSSION

We detected high proportions of realized climatic niche shifts in invaded ranges

For Peer Review

for both non-native reptile and amphibian species at a global scale. The climatic niche shifts were more likely if a non-native species had a smaller native range, was introduced into a new range earlier or invaded lower latitudinal areas than its native range. A recent study of the drivers of climatic niche dynamics (niche expansion, niche overlap and niche unfilling) for European non-native birds (Strubbe *et al.*, 2013) has found that niche overlap was lower for more recently introduced species and that niche unfilling was lower for species that were introduced earlier and more often. The latter study did not detect any significant variables to explain the niche shifts, due partly to the low proportion of niche shifts detected for birds (Strubbe *et al.*, 2013). As far as we are aware, our study is the first to identify factors influencing realized climatic niche shifts between invaded and native ranges for non-native species.

The results of this study highlight the importance of understanding the factors that influence realized niche shifts that occur between the invaded and native ranges of species. Previous studies often offered, as potential explanations of the observed niche shifts, that both evolutionary and ecological changes may take place: changes of the fundamental niche limits and associated evolutionary changes, and changes of the factors – dispersal and biotic interactions – that are constraining the occupancy of the fundamental niche in the native range (Fitzpatrick *et al.*, 2007; Broennimann & Guisan, 2008; Beaumont *et al.*, 2009). It is probable, however, that observed climatic niche shifts for non-native herpetofauna result more from ecological than from evolutionary processes. The effects of residence time and equatorward expansion in invaded range and native range size on the niche shifts that we identified tend to

For Peer Review

confirm the general hypothesis that climatic niche shifts for non-native species result from changes in dispersal limitations and biotic interactions between native and invaded ranges and from the introduction history (Pearman *et al.*, 2008; Alexander & Edwards, 2010). The high proportion of realized niche shifts in the invaded range implies that many non-native herpetofauna species only occupy, in the native and/or invaded range, a part of the environment that is potentially suitable due to dispersal limitations and changes in biological interactions.

The proportion of niche shifts for non-native herpetofauna was higher than those for non-native plants and birds (Petitpierre *et al.*, 2012; Strubbe *et al.*, 2013) even if island endemics were excluded from the analysis (Tables S7D-F). Clearly, the higher proportion of island endemics was partly responsible for differences between our study and the two other studies cited. In addition, the introduction histories showed large differences among the three taxa (median introduction date for non-native herpetofauna, 1917; median introduction date for plants, 1750; and median introduction date for birds, 1963). The earlier introduction dates for herpetofauna than for bird species might partly explain the higher proportion of niche shifts for the former compared to the latter, but this difference in introduction histories does not explain the difference in niche shifts between the herpetofauna and plants. However, much larger native range sizes may explain the lower niche shifts in plants (mean = 35,201, SD = 25,143 10 arc-minute grid cells) compared with the herpetofauna species studied here. Finally, differences in the variables used in the studies might be partly responsible for the differences in niche shifts. All three studies included

For Peer Review

important (eight) climate variables to characterize the realized climatic niches of the studied taxa, but the study on birds also included the human footprint (Strubbe *et al.*, 2013).

Island endemics were more likely not to be at equilibrium with climate due to the ocean as a barrier to dispersal and therefore showed a higher proportion of niche shifts. The higher proportion of niche shifts for invaded ranges in Oceania than in the combination of the other six realms might be due to the difference in native range size. The species introduced to Oceania had a smaller range size than those introduced to the other realms (χ^2 = 4.016, *df* = 1, *P* = 0.045), but no differences in residence time (χ^2 = 0.039, *P* = 0.843) or equatorward expansion were detected for introduced species between Oceania and the other realms (χ^2 = 0.017, *P* = 0.896). As a result, species introduced to Oceania may be more likely to exhibit niche shifts than those introduced to other realms.

We found weak evidence for the effects of island endemic identity on niche expansion in the model averaging analysis (Table 2). This may have arisen because island endemic identity was negatively correlated with native range size (Table S9). Due to sea barriers, island endemics had a smaller native range size than others (χ^2 = 15.379, *df* = 1, *P* < 0.001), which might explain most of the variation in niche shifts resulting from island endemic identity.

The minor influence of the number of introduction events on the niche shifts may result from two factors. First, for certain species, the admixtures resulting from multiple introductions of different local populations from the native range might

For Peer Review

result in offspring with decreased levels of viability or fertility due to the incompatibility of genomes from divergent lineages (Johnson, 2010). Because the hybrids with decreased fitness would be poorly adapted to climate habitats in the hybrid zone compared with their parents, rapid fundamental niche shifts for such hybrids could not occur (also supporting our hypothesis that changes result more from ecological than from evolutionary factors). Second, certain species might have been introduced multiple times into geographically disconnected areas of a realm. Because the introduced populations did not mix, the genetic variation in the populations could not increase, resulting in no effect of the number of introduction events on the niche shifts. Consistent with the results of a previous study of birds (Strubbe *et al.*, 2013), the invaded range size (number of grid cells) was correlated with the number of introduction events in this study (r = 0.35, df = 100, P < 0.001), suggesting that multiple introductions resulted in a larger invaded range.

One potential explanation for the weak effect of poleward expansion was that the high-latitude boundary of native range for the species with poleward expansion might be less limited by climate. As a result, increasing poleward expansion in invaded ranges might not increase climatic niche expansion. It is known that dispersal (Araújo & Pearson, 2005; Araújo *et al.*, 2008; Baselga *et al*., 2012) or the conservatism of cold physiological tolerance (Olalla-Tarraga *et al*., 2011) determines the high-latitude range boundary for amphibians and reptiles. The distributions of restricted-range species were more likely to be limited by intrinsic dispersal limitation (low dispersal abilities) but less by climate than those of widespread species (Araújo & Pearson,

For Peer Review

2005; Baselga *et al*., 2012). Poleward expansion in invaded ranges for non-native herpetofauna was negatively correlated with their native range sizes (Table S9), indicating that the species with smaller range size had increased poleward expansion in their invaded range. Climatic factors, including cold tolerance, might pose fewer constraints on the higher-latitude boundary of the native range for these species due to their smaller native range size.

Our results may have important implications for ENM-based predictions of the risk of biological invasions by reptiles and amphibians. 'Climate matching' between invaded and native ranges and propagule pressures are commonly used to guide risk assessments for initial introductions (Lockwood *et al.*, 2005; Bomford *et al.*, 2009; Guisan *et al*. 2013). Our results suggested that 'climate matching' based on ENMs should be applied with caution for species with small native ranges or those that are introduced into lower-latitude areas compared with the native range because realized climatic niche shifts are more likely for these species. In such cases, 'climate matching' might underestimate the risk of establishment of the non-native species. Because both the fundamental niche and the realized niche are susceptible to shift with increasing residence time for established non-native species, it is advisable to rerun ENMs in a timely manner based on comprehensive occurrence data from the entire distribution, including both the native range and the new invaded ranges. These updated analyses may produce better predictions of the risk represented by such species (Broennimann & Guisan, 2008; Beaumont *et al.*, 2009; Gallagher *et al.*, 2010).

For Peer Review

In conclusion, we found that climatic niche shifts were pervasive among non-native herpetofauna, a pattern that was strikingly different from those of Holarctic non-native plants and European non-native birds. Therefore, 'climate matching' as a guide for risk assessments of initial introductions was not warranted for the non-native herpetofauna. We identified residence time, equatorward expansion in the invaded range and native range size as factors influencing climate niche shifts. These results partly explain how the debate on niche conservation between invaded and native ranges for non-native species can arise (e.g., in terms of differences among the non-native plants, birds and herpetofauna in this study). Our results provide insights into the climatic niche dynamics for both native and invaded ranges. This understanding may facilitate better ENM-based predictions of the risks of biological invasions.

ACKNOWLEDGEMENTS

We thank Zetian Liu and Shaofei Yan for assisting with the collection of data, F. Kraus for helpful suggestions on data collection, and David Currie, Diniz-Filho, José Alexandre and four anonymous reviewers for their valuable comments on the manuscript. This research was supported by grants from National Science Foundation of China (31370545 and 31200416) and a grant from The Ministry of Science and Technology of China (2013FY110300).

REFERENCES

For Peer Review

Alexander, J.M. & Edwards, P.J. (2010) Limits to the niche and range margins of alien species. *Oikos*, **119**, 1377-1386.

Araújo, M.B. & Pearson, R.G. (2005) Equilibrium of species' distributions with climate. *Ecography*, **28**, 693-695.

Araújo, M.B. & Peterson, A.T. (2012) Uses and misuses of bioclimatic envelope modeling. *Ecology*, **93**, 1527-1539.

Araújo, M.B., Nogués-Bravo, D., Diniz-Filho, J.A.F., Haywood, A.M., Valdes, P.J. & Rahbek, C. (2008) Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography*, **31**, 8-15.

Aragón, P., Lobo, J.M., Olalla-Tárraga, M.Á. & Rodríguez, M.Á. (2010) The contribution of contemporary climate to ectothermic and endothermic vertebrate distributions in a glacial refuge. *Global Ecology and Biogeography*, **19**, 40-49.

Barve, N., Barve, V., Jimenez-Valverde, A., Lira-Noriega, A., Maher, S.P., Peterson, A.T., Soberón, J. & Villalobos, F. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **222**, 1810-1819.

Baselga, A., Lobo, J.M., Svenning, J.C. & Araújo, M.B. (2012) Global patterns in the shape of species geographical ranges reveal range determinants. *Journal of Biogeography*, **39**, 760-771.

Beaumont, L.J., Gallagher, R.V., Thuiller, W., Downey, P.O., Leishman, M.R. & Hughes, L. (2009) Different climatic envelopes among invasive populations

For Peer Review

may lead to underestimations of current and future biological invasions. *Diversity and Distributions*, **15**, 409-420.

Beyer, H. (2010) Hawth's analysis tools for ArcGIS. 2004. *Available at the following web site: www. spatialecology.com/htools*.

Bomford, M., Kraus, F., Barry, S.C. & Lawrence, E. (2009) Predicting establishment success for alien reptiles and amphibians: a role for climate matching. *Biological Invasions*, **11**, 713-724.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, **4**, 585-589.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C. & Zimmermann, N.E. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481-497.

Burnham, K. & Anderson, D. (2002) *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Verlag.

Darwin, C. (1859) *On the origin of species London*. UK: John Murray.

Fitzpatrick, M. & Hargrove, W. (2009) The projection of species distribution models and the problem of non-analog climate. *Biodiversity and Conservation*, **18**, 2255-2261.

Fitzpatrick, M.C., Weltzin, J.F., Sanders, N.J. & Dunn, R.R. (2007) The biogeography of prediction error: why does the introduced range of the fire ant over-predict

For Peer Review

its native range? *Global Ecology and Biogeography*, **16**, 24-33.

Gallagher, R.V., Beaumont, L.J., Hughes, L. & Leishman, M.R. (2010) Evidence for climatic niche and biome shifts between native and novel ranges in plant species introduced to Australia. *Journal of Ecology*, **98**, 790-799.

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

Guisan, A., Petitpierre, B., Broennimann, O., Kueffer, C., Randin, C. & Daehler, C. (2012) Response to Comment on "Climatic Niche Shifts Are Rare Among Terrestrial Plant Invaders". *Science*, **338**, 193.

Guisan, A., Tingley, R., Baumgartner, J.B., Naujokaitis-Lewis, I., Sutcliffe, P.R., Tulloch, A.I.T., Regan, T.J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T.G., Rhodes, J.R., Maggini, R., Setterfield, S.A., Elith, J., Schwartz, M.W., Wintle, B.A., Broennimann, O., Austin, M., Austin, M., Kearney, M.R., Possingham, H.P. & Buckley, Y.M. (2013) Predicting species distributions for conservation decisions. *Ecology Letters*, **16**, 1424-1435.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965-1978.

Johnson, N.A. (2010) Hybrid incompatibility genes: remnants of a genomic battlefield? *Trends in Genetics*, **26**, 317-325.

Kolbe, J.J., Glor, R.E., Schettino, L.R., Lara, A.C., Larson, A. & Losos, J.B. (2004)

For Peer Review

Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177-181.

Kraus, F. (2009) *Alien reptiles and amphibians: a scientific compendium and analysis*. Springer Verlag.

Liu, X., Rohr, J.R. & Li, Y. (2013) Climate, vegetation, introduced hosts and trade shape a global wildlife pandemic. *Proceedings of the Royal Society B: Biological Sciences*, **280,** 20122506.

Lockwood, J., Cassey, P. & Blackburn, T. (2005) The role of propagule pressure in explaining species invasions. *Trends in Ecology & Evolution*, **20**, 223-228.

Losos, J.B. (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, **11**, 995-1003.

Normand, S., Treier, U.A., Randin, C., Vittoz, P., Guisan, A. & Svenning, J.-C. (2009) Importance of abiotic stress as a range-limit determinant for European plants: insights from species responses to climatic gradients. *Global Ecology and Biogeography*, 18, 437-449.

Olalla-Tarraga, M.A., McInnes, L., Bini, L.M., Diniz-Filho, J.A.F., Fritz, S.A., Hawkins, B.A., Hortal, J., Orme, C.D.L., Rahbek, C., Rodriguez, M.A. & Purvis, A. (2011) Climatic niche conservatism and the evolutionary dynamics in species range boundaries: global congruence across mammals and amphibians. *Journal of Biogeography*, **38**, 2237-2247.

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N.,

For Peer Review

Underwood, E.C., D'amico, J.A., Itoua, I., Strand, H.E. & Morrison, J.C. (2001) Terrestrial ecoregions of the world: a new map of life on earth. *BioScience*, **51**, 933-938.

Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008) Niche dynamics in space and time. *Trends in Ecology & Evolution*, **23**, 149-158.

Pearman, P.B., Lavergne, S., Roquet, C., Wüest, R., Zimmermann, N.E. & Thuiller, W. (2013) Phylogenetic patterns of climatic, habitat and trophic niches in a European avian assemblage. *Global Ecology and Biogeography*, in press.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, **38**, 817-827.

Peterson, A.T. & Nakazawa, Y. (2008) Environmental data sets matter in ecological niche modelling: an example with *Solenopsis invicta* and *Solenopsis richteri*. *Global Ecology and Biogeography*, **17**, 135-144.

Peterson, A.T., Soberón, J., Pearson, R.G., Anderson, R.P., Martinez-Meyer, E., Nakamura, M. &Araújo, M.B. (2011) *Ecological Niches and Geographic Distributions*. Princeton University Press.

Petitpierre, B., Kueffer, C., Broennimann, O., Randin, C., Daehler, C. & Guisan, A. (2012) Climatic niche shifts are rare among terrestrial plant invaders. *Science*, **335**, 1344-1348.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

For Peer Review

Sax, D.F., Stachowicz, J.J., Brown, J.H., Bruno, J.F., Dawson, M.N., Gaines, S.D., Grosberg, R.K., Hastings, A., Holt, R.D., Mayfield, M.M., O'Connor, M.I. & Rice, W.R. (2007) Ecological and evolutionary insights from species invasions. *Trends in Ecology & Evolution*, **22**, 465-471.

Schulte, U., Hochkirch, A., Lötters, S., Rödder, D., Schweiger, S., Weimann, T. & Veith, M. (2012) Cryptic niche conservatism among evolutionary lineages of an invasive lizard. *Global Ecology and Biogeography*, **21**, 198-211.

Soberón, J. & Nakamura, M. (2009) Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, **106**, 19644-19650.

Soberón, J. & Peterson, A.T. (2011) Ecological niche shifts and environmental space anisotropy: a cautionary note. *Revista Mexicana de Biodiversidad*, **82**

Strubbe, D., Broennimann, O., Chiron, F. & Matthysen, E. (2013) Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography*, **22**, 962-970.

Tingley, R., Phillips, B.L. & Shine, R. (2011) Establishment success of introduced amphibians increases in the presence of congeneric species. *The American Naturalist*, **177**, 382-388.

Van Wilgen, N.J. & Richardson, D.M. (2012) The Roles of Climate, Phylogenetic Relatedness, Introduction Effort, and Reproductive Traits in the Establishment of Non-Native Reptiles and Amphibians. *Conservation Biology*, **26**, 267-277.

For Peer Review

Vitt, L.J. & Caldwell, J.P. (2009) *Herpetology*. Academic Press.

Warren, D.L., Glor, R.E. & Turelli, M. (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868-2883.

Whitton, F.J.S., Purvis, A., Orme, C.D.L. & Olalla-Tarraga, M.A. (2012) Understanding global patterns in amphibian geographic range size: does Rapoport rule? *Global Ecology and Biogeography*, **21**, 179-190.

Wiens, J.J., Ackerly, D.D., Allen, A.P., Anacker, B.L., Buckley, L.B., Cornell, H.V., Damschen, E.I., Davies, T.J., Grytnes, J.A., Harrison, S.P., Hawkins, B.A., Holt, R.D., McCain, C.M. & Stephens, P.R. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310-1324.

Williamson, M., Dehnen-Schmutz, K., Kühn, I., Hill, M., Klotz, S., Milbau, A., Stout, J. & Pyšek, P. (2009) The distribution of range sizes of native and alien plants in four European countries and the effects of residence time. *Diversity and Distributions*, **15**, 158-166.

Wilson, J.R.U., Richardson, D.M., Rouget, M., Procheş, Ş., Amis, M.A., Henderson, L. & Thuiller, W. (2007) Residence time and potential range: crucial considerations in modelling plant invasions. *Diversity and Distributions*, **13**, 11-22.

For Peer Review

BIOSKETCHES

Dr. **Yiming Li** is a professor of animal ecology and conservation biology. His research focuses on the conservation and ecology of amphibians, and macroecology of vertebrates. Dr. **Xuan Liu** is an assistant professor interested in the macroecology, biogeography and conservation biology with specific focus on impacts of global change factors on biodiversity. **Xianping Li** is a PhD candidate investigating the climatic niche conservatism of non-native species and their determinants.

Author contribution: YL and XLIU designed the study; XL and XLIU collected data; XLIU, XL, YL, BP and AG analyzed data; YL, XLIU, XL, BP and AG wrote the manuscript.

For Peer Review

Tables

Table 1. The six best models (i.e., $\Delta AIC_c \leq 2$) of the factors influencing niche shifts for 101 invaded ranges of 71 reptile and amphibian invaders worldwide. The models are based on a binomial error structure and a logit link function with climatic niche shifts (≥10% expansion = 1, others = 0) as the response variable, a combination of six factors as explanatory variables and the native and invaded biogeographic realms as random variables. The native realm or invaded realm showed no effects on niche shifts as random variables (for details, see Table S5).

Variables	Model rank					
	1	2	3	4	5	6
Native Range Size (grid cells)	•	•	•	•	•	•
Residence Time (years)	•	•	•	•	•	•
Equatorward expansion in invaded range (degrees)	•	•	•	•	•	•
Poleward expansion in invaded ranges (degrees)	•		•		•	
Number of introduction events		•			•	•
Island Endemic identity (binary variable)			•			•
ΔAIC	0	0.19	0.69	1.17	1.40	1.43
AIC_c	84.39	84.58	85.07	85.56	85.78	85.81
W_i	0.21	0.19	0.15	0.12	0.10	0.10
R^2	0.66	0.66	0.67	0.64	0.67	0.67

•: indicates that a variable is included in the model

ΔAIC: the difference between each model and the highest ranked model

AIC_c: Akaike's information criterion adjusted for small sample sizes

W_i (Akaike weights): the relative likelihood of a model given the particular set of best models being considered

R^2: Likelihood-ratio based pseudo-R-squared

For Peer Review

Table 2. Summary of multiple-model inference for 101 invaded ranges of 71 non-native amphibian and reptile species at a global scale. The model averaging was based on generalized linear mixed models (2^6-1 models) with a binomial error structure and a logit link function with climatic niche shifts (≥10% expansion = 1, others = 0) as the response variable, a combination of six factors as explanatory variables and the native and invaded biogeographic realms as random variables. The effect of the native or invaded realm was significant only in three (one) of 63 models (for details, see Table S5).

Explanatory variables	**Relative importance**	**Parameter estimate (SE)**
Native Range Size(grid cells)	1.00	-0.007 (0.002)
Residence Time (years)	1.00	0.020 (0.008)
Equatorward expansion in invaded range (degrees)	1.00	3.650 (3.520)
Poleward expansion in invaded ranges (degrees)	0.53	0.126 (0.084)
Number of introduction events	0.46	0.021 (0.034)
Island Endemic identity (binary variable)	0.38	1.333 (1.129)

For Peer Review

Figure legends

Fig. 1. Examples of climatic niche dynamics between native and invaded ranges: (a) the American bullfrog (*Lithobates catesbeianus*) introduced within the native biogeographic realm (Nearctic) and (b) the red-eared slider (*Trachemys scripta elegans*) introduced from the Nearctic to the Palearctic. The solid and dashed contour lines indicate 100% and 75%, respectively, of the available (background) environment (green: the native climatic background; red: the invasive background). Green area: unfilling; blue area: stability; and red area: expansion. The solid arrows (dashed arrows) represent the change in the center of the species niche (climatic space) between the native and invaded ranges (the dashed line in figure (a) is superimposed, as the introduction event occurred in the same biogeographic realm).

Fig. 2. Proportions of invaded ranges or species showing a climatic niche shift (≥10% expansion) for non-native herpetofauna: (a) at the invaded range level (left) and at the species level (right); (b) invaded ranges within the native biogeographic realm and those in different biogeographic realms (left), island endemics and other species (right).

Fig. 3. Proportions of invaded ranges showing a climatic niche shift (≥10% expansion) for non-native herpetofauna across each biogeographic realm. The denominator in a pie slice represents the sample size of invaded ranges and the numerator represents the number of invaded ranges showing a climate niche shift, respectively. The size of

For Peer Review

the slice in the right half of a pie represents the relative sample size between two taxa.

Supporting Information legends

Table S1 Databases and literature used for occurrence data.

Table S2 Niche change indices based on different extents and resolutions.

Table S3 Ranking of the eight climatic variables by PCA.

Table S4 Significance test for phylogenetic signal.

Table S5 Detailed results of multiple-model inference.

Table S6 Data on niche expansion and six predictor variables.

Table S7 Comparisons of niche dynamics between extents, resolutions and taxa.

Table S8 Multiple-model inference based on the different extent and sample.

Table S9 Spearman rank correlation coefficients among six predictors.

Text S1 Methods for the database of species list.

For Peer Review

Fig. S1 Correlations between climatic variables and first two components of PCA.

Fig. S2 Sensitivity analysis to the marginal climates.

Fig. S3 Scheme of the full data analysis procedure.

Fig. S4 Climatic niche dynamics for non-native herpetofauna.

For Peer Review

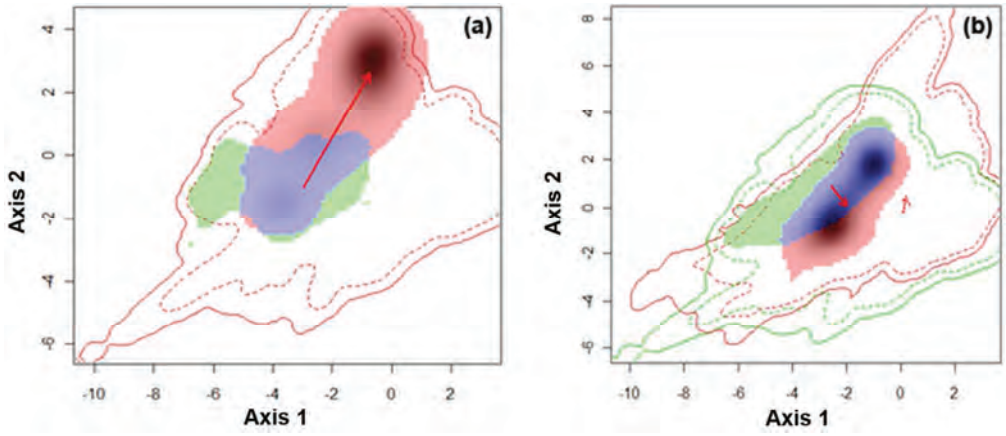


Fig. 1. Examples of climatic niche dynamics between native and invaded ranges: (a) the American bullfrog (Lithobates catesbeianus) introduced within the native biogeographic realm (Nearctic) and (b) the red-eared slider (Trachemys scripta elegans) introduced from the Nearctic to the Palearctic. The solid and dashed contour lines indicate 100% and 75%, respectively, of the available (background) environment (green: the native climatic background; red: the invasive background). Green area: unfilling; blue area: stability; and red area: expansion. The solid arrows (dashed arrows) represent the change in the center of the species niche (climatic space) between the native and invaded ranges (the dashed line in figure (a) is superimposed, as the introduction event occurred in the same biogeographic realm).
96x41mm (300 x 300 DPI)

For Review

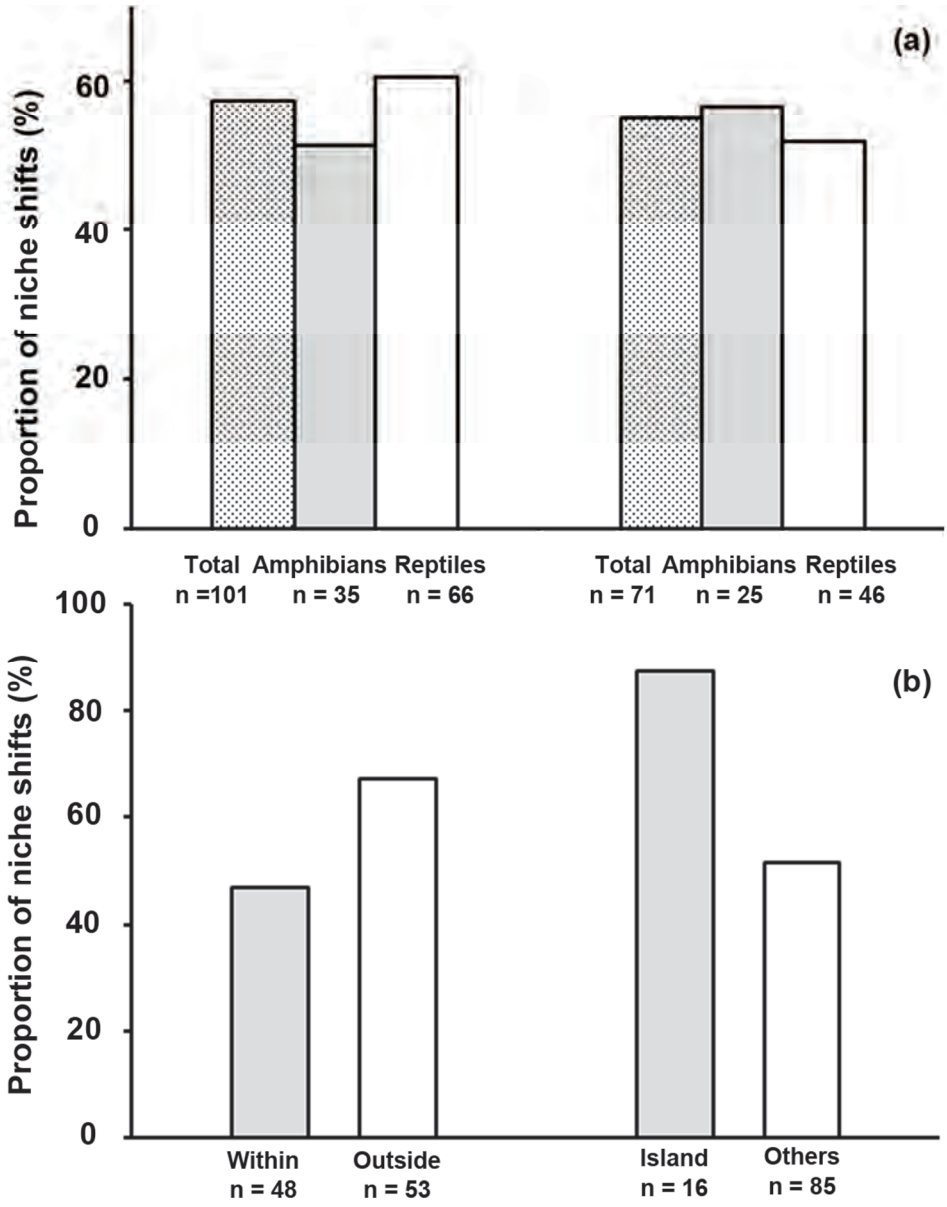


Fig. 2. Proportions of invaded ranges or species showing a climatic niche shift (≥10% expansion) for non-native herpetofauna: (a) at the invaded range level (left) and at the species level (right); (b) invaded ranges within the native biogeographic realm and those in different biogeographic realms (left), island endemics and other species (right).
99x123mm (300 x 300 DPI)

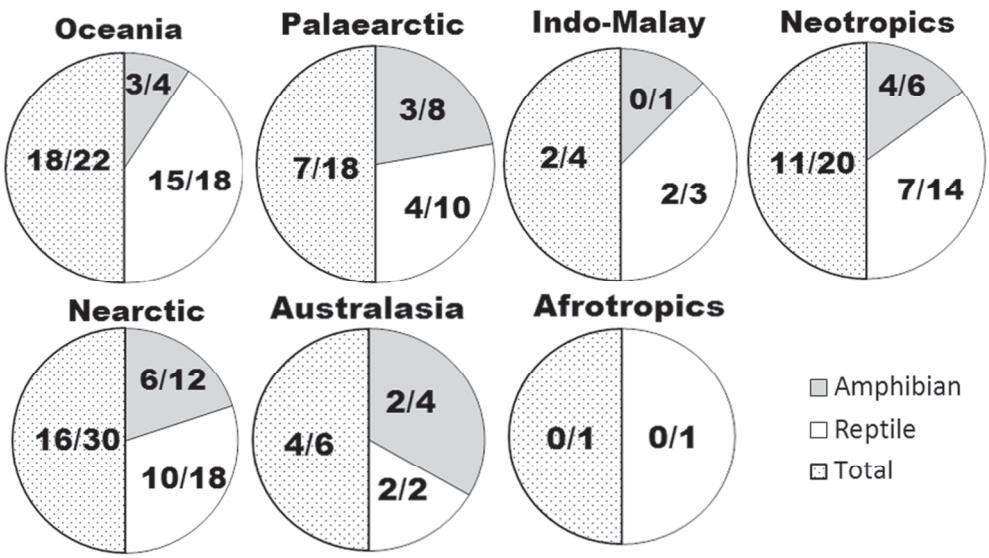


Fig. 3. Proportions of invaded ranges showing a climatic niche shift (≥10% expansion) for non-native herpetofauna across each biogeographic realm. The denominator in a pie slice represents the sample size of invaded ranges and the numerator represents the number of invaded ranges showing a climate niche shift, respectively. The size of the slice in the right half of a pie represents the relative sample size between two taxa.
99x55mm (300 x 300 DPI)

Review

Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2013) **22**, 302–317

A Journal of Macroecology

Global Ecology and Biogeography

SPECIAL ISSUE

Building the niche through time: using 13,000 years of data to predict the effects of climate change on three tree species in Europe

L. Maiorano[1]*, R. Cheddadi[2], N. E. Zimmermann[3], L. Pellissier[1], B. Petitpierre[1], J. Pottier[1], H. Laborde[2], B. I. Hurdu[4,5], P. B. Pearman[3], A. Psomas[3], J. S. Singarayer[6], O. Broennimann[1], P. Vittoz[1], A. Dubuis[1], M. E. Edwards[7], H. A. Binney[7] and A. Guisan[1]

[1]*Department of Ecology and Evolution, University of Lausanne, Biophore Building, CH-1015 Lausanne, Switzerland,* [2]*CNRS, University of Montpellier-2, Institut des Sciences de l'Évolution, CC 061, Place Eugène-Bataillon, 34095 Montpellier Cedex 05, France,* [3]*Swiss Federal Research Institute WSL, Zuercherstrasse 111, CH-8903 Birmensdorf, Switzerland,* [4]*Faculty of Biology and Geology, Babeş-Bolyai University, 400015 Cluj-Napoca, Romania,* [5]*Institute of Biological Research, 400015 Cluj-Napoca, Romania,* [6]*Department of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK,* [7]*Geography and Environment, University of Southampton, Highfield, Southampton SO17 1BJ, UK*

*Correspondence: Luigi Maiorano, Department of Ecology and Evolution, University of Lausanne, Biophore Building, CH-1015 Lausanne, Switzerland.
E-mail: luigi.maiorano@unil.ch

ABSTRACT

Aim Species distribution models (SDMs) based on current species ranges underestimate the potential distribution when projected in time and/or space. A multi-temporal model calibration approach has been suggested as an alternative, and we evaluate this using 13,000 years of data.

Location Europe.

Methods We used fossil-based records of presence for *Picea abies*, *Abies alba* and *Fagus sylvatica* and six climatic variables for the period 13,000 to 1000 yr BP. To measure the contribution of each 1000-year time step to the total niche of each species (the niche measured by pooling all the data), we employed a principal components analysis (PCA) calibrated with data over the entire range of possible climates. Then we projected both the total niche and the partial niches from single time frames into the PCA space, and tested if the partial niches were more similar to the total niche than random. Using an ensemble forecasting approach, we calibrated SDMs for each time frame and for the pooled database. We projected each model to current climate and evaluated the results against current pollen data. We also projected all models into the future.

Results Niche similarity between the partial and the total-SDMs was almost always statistically significant and increased through time. SDMs calibrated from single time frames gave different results when projected to current climate, providing evidence of a change in the species realized niches through time. Moreover, they predicted limited climate suitability when compared with the total-SDMs. The same results were obtained when projected to future climates.

Main conclusions The realized climatic niche of species differed for current and future climates when SDMs were calibrated considering different past climates. Building the niche as an ensemble through time represents a way forward to a better understanding of a species' range and its ecology in a changing climate.

Keywords

Climate change, Europe, forecasting, fundamental niche, Holocene, realized niche, species distribution models.

DOI: 10.1111/j.1466-8238.2012.00767.x

© 2012 Blackwell Publishing Ltd http://wileyonlinelibrary.com/journal/geb

INTRODUCTION

The geographical distribution of a species depends on a wide variety of biotic, abiotic and historical processes operating at different spatial and temporal scales (Guisan & Thuiller, 2005), with climate playing a prominent role especially at larger spatial scales. Bioclimatic species distribution models (SDMs) represent one of the most important available tools to assess the likely impact of global warming, and their use is growing rapidly in many areas of ecology and conservation (Nogués-Bravo *et al.*, 2008; Maiorano *et al.*, 2011).

Probably the most important underlying ecological concept for SDMs is a species' ecological niche. Hutchinson (1957) defined the fundamental niche as comprising those environmental abiotic conditions that allow a species to maintain sustainable population growth, and the realized niche as the part of the fundamental niche that can be occupied when competition or other forms of biotic interactions are included. Although the aim when using SDMs to project species distribution in time and space is to represent the fundamental niche as closely as possible (Colwell & Rangel, 2009), SDMs are in fact based on the empirical relationships between observed species distributions and environmental variables (Guisan & Thuiller, 2005). Thus, SDMs essentially characterize bioclimatic envelopes based on the realized niche, since observed species distributions are constrained by non-climatic factors, including biotic interactions (Meier *et al.*, 2011a).

When projecting SDMs in space and/or in time a number of issues may influence the results, ranging from theoretical assumptions (Guisan & Thuiller, 2005), to model techniques (Hijmans & Graham, 2006), to projections in non-analogue climates (Williams & Jackson, 2007). However, the two most important assumptions are related to niche stability and to the equilibrium between species distribution and climate (Araújo & Pearson, 2005; Pearman *et al.*, 2008a).

Niche stability assumes that the range of climatic conditions that allows a single species to persist in its geographical range through time does not change, and that no significant evolutionary and/or ecological changes affected a species' niche, even when environmental conditions change in time or space. However, evidence suggests that niche shifts may have occurred, at least for some species (Pearman *et al.*, 2008b), possibly as the result of genetic variation in traits related to climate (Eberling *et al.*, 2008), thus expressing a change in the fundamental niche, or possibly linked to shifts in competition among species over time (Nogués-Bravo, 2009), expressing a change in the realized niche.

SDMs further assume equilibrium between species distribution and climate (Araújo & Pearson, 2005). Yet the species may not colonize potentially suitable areas due to dispersal limitations and/or to biotic interactions. Currently, for example, many tree species in Europe do not seem to be fully in equilibrium with climate (Normand *et al.*, 2011), probably as a consequence of post-glacial dispersal limitations (Svenning *et al.*, 2008).

Even if all theoretical assumptions hold, SDMs trained from data on current species ranges risk underestimating the potential suitable environment and geographical range for species when projected to new areas or time periods (Dormann, 2007) because of dispersal limitation, biotic interaction or other constraining factors. In fact, even in the absence of biotic and/or dispersal limitations, SDMs calibrated considering a single time period will only provide a partial view of the fundamental niche, as it is likely that current conditions do not encompass the complete climate space that a species could potentially occupy (Varela *et al.*, 2009).

Different approaches have been suggested to overcome the problems due to deviations from the assumptions of niche stability and equilibrium. The classic solutions have been to consider only species that are at or near equilibrium with climate (Nogués-Bravo, 2009), and to sample as much as possible of the global distribution of a species (Broennimann & Guisan, 2008). However, Varela *et al.* (2009) demonstrated that even a combination of these two solutions is not sufficient and can actually result in limited model transferability in time. A more direct solution is to measure the physiological tolerance of the species to climatic conditions directly (Kearney & Porter, 2009), but application is limited to a few well-studied species.

A new approach, based on multi-temporal model calibration, has recently been suggested as a viable alternative (Nogués-Bravo, 2009; Varela *et al.*, 2009). This approach was first suggested by Broennimann & Guisan (2008) who calibrated a SDM for *Centaurea stoebe* by pooling presence data from the native and invaded ranges in Europe and North America. They found that projections from pooled SDMs showed a better fit and reduced uncertainty under future projections. Nogués-Bravo *et al.* (2008) transposed the same approach to time when calibrating models for the woolly mammoth. Here, the climatic niche of the mammoth was calibrated using the geographical location of dated fossil remains belonging to three different time periods. The projection of this multi-temporal climatic niche, evaluated against independent fossil data, accurately modelled the distribution of the species. The theoretical background behind this approach is simple: a hypothetical species X is shifting its realized climatic niche through time; assuming no change in the fundamental niche of X, a SDM obtained through multi-temporal calibration would provide a better approximation of the fundamental niche (Nogués-Bravo, 2009) or at least a more complete realized niche covering a broader range of possible climate responses than is observable with one single time period.

Using palaeoclimate simulations and fossil records of three tree species (i.e. *Picea abies* (L.) H. Karst., *Abies alba* Mill. and *Fagus sylvatica* L.), we adopted an ensemble forecasting approach to build a multi-temporal climate envelope considering Europe for the past 13,000 years. We combine SDMs developed from past climate and fossil records to project them to future climates and quantify the contribution of each time frame to the total species niche to demonstrate the value of the multi-temporal approach. Specifically, we build SDMs related to each single time frame (partial-SDM) and to all time frames at once (total-SDM), and we project all models to the current climate evaluating the projections against the current pollen

samples. We also project the total-SDM into the future in order to obtain improved predictions of the potential response of species to climate change.

MATERIALS AND METHODS

Species data

Records of presence for spruce (*P. abies*), silver fir (*A. alba*) and beech (*F. sylvatica*) from 13,000 to 1000 yr BP were obtained from two main sources: (1) the European Pollen Database (EPD; http://www.europeanpollendatabase.net) and (2) a review of published data including both pollen and macrofossils (Terhürne-Berson *et al.*, 2004; van der Knaap *et al.*, 2005; Latalowa & van der Knaap 2006; Magri *et al.*, 2006). All locations with the associated pollen records from the EPD were associated to one or more single-millennium time frames (± 500 years) using calibrated ^{14}C dates. The number of occurrences varies in each time frame, for a total of 1078 points for the silver fir, 801 for beech and 2217 for spruce (Table 1). Pollen percentages obtained from the EPD were registered as species presences using a 1% threshold as recommended by van der Knaap *et al.* (2005). The same threshold was applied to modern pollen samples available from the EPD (core tops only) and from Guiot *et al.* (1996), for a total of 341 points for the silver fir, 633 for beech and 668 for spruce (Fig. S1 in Supporting Information). The current species distribution (presence/absence records) for the three species was obtained from the Atlas Florae Europaea (Jalas & Suominen, 1972–1999).

Current climate and climate projections

We considered a set of six bioclimatic variables assumed to be important for the ecology and physiology of the species considered: total annual precipitation, summer (June, July and August) precipitation, winter (December, January and February) precipitation, annual mean temperature, summer mean temperature and winter mean temperature.

Table 1 Number of points of presence available for each species in each time frame (number of macrofossils in parenthesis).

Years before present	*Abies alba*	*Fagus sylvatica*	*Picea abies*
13,000	6 (0)	–	17 (0)
12,000	15 (1)	3	31 (7)
11,000	13 (1)	5	52 (9)
10,000	11 (6)	16 (3)	72 (18)
9000	24 (11)	25 (0)	98 (20)
8000	52 (10)	40 (3)	127 (42)
7000	90 (8)	73 (5)	153 (49)
6000	106 (11)	110 (4)	202 (57)
5000	122 (0)	136 (4)	253 (89)
4000	125 (0)	185 (0)	286 (81)
3000	125 (0)	208 (1)	301 (79)
2000	214 (0)	–	310 (47)
1000	175 (0)	–	315 (45)

Data on current climate (averaged from 1950 to 2000) and future climate projections (average from 2070 to 2100) were obtained from the Climatic Research Unit (Mitchell *et al.*, 2004). Simulations of past climate were obtained for the same time period considered for the fossil pollen data from a global ocean–atmosphere climate model with a temporal resolution of 1000 years and a spatial resolution of 3.75° × 2.5° (Singarayer & Valdes, 2010) based on the Hadley Centre climate model (HadCM3). We first calculated climate anomalies by contrasting the past monthly temperature and precipitation values against the pre-industrial climate as obtained from the same simulation (Singarayer & Valdes, 2010). Then, given the existing offset of several degrees in temperature between the pre-industrial simulation and the current climate, we calculated a second set of climate anomalies by contrasting climate means for 1901–20 against the current climate (both obtained from Mitchell *et al.*, 2004), assuming that the former provides an approximation closer to the real pre-industrial climate than the available simulations. The combination of these two anomalies allowed us to correct for the bias inherent in each climate model output. This avoids transition inconsistencies when combining measured (current) and simulated (past/future) climates, because simulated climate runs do not match exactly with measured climate averages. Anomalies were calculated in both cases as absolute temperature difference (Δ°C) and relative precipitation differences (% change).

Using bilinear resampling, we downscaled the first set of anomalies (past climate against pre-industrial) to 0.5° of spatial resolution, corresponding to the spatial resolution of the current climate dataset. Then, we applied the anomaly corrections to obtain monthly maps of past temperature and precipitation at a 0.5° spatial resolution. Finally, we calculated all derived climate maps mentioned above.

To be consistent with the climate model considered in the past, for future projections we considered the same HadCM3 climate change model and two of the IPCC's (Intergovernmental Panel on Climate Change) climatic scenarios, the B1 scenario, describing a world with reduced use of natural resources and the use of clean and resource-efficient technologies, and the A2 scenario, where the rate of emission of greenhouse gases continues to increase.

Niche expansion through time

To measure the contribution of each time step (i.e. data from the available 1000-year frames) to the total niche of the species (i.e. the niche as measured considering all available points), we considered the framework proposed by Broennimann *et al.* (2012): niche similarity is measured in a gridded and smoothed principal components analysis (PCA) environmental space, taking in account environmental availability (i.e. the values of the climate variables over the study area for each time frame).

We represented the total environmental space by means of the first two axes of a PCA calibrated over the entire range of

possible environments in the study area (i.e. considering all climate variables as measured for 2000 random points drawn for each time frame). Then, we pooled all pollen data through time, measured the values of the climate variables corresponding to pollen locations and time frames, and projected them into the PCA space to delineate the total niche space for each species. We followed the same process for each time frame separately, projecting all the points of presence for each millennium in turn into the total PCA space in order to delineate the partial niches for each species per time frame. Then we compared the partial niches with the total niche following Broennimann *et al.* (2012) and tested if the partial niches were more similar to the total niche than random using Schoener's *D* (an index of overlap). To account for prevalence effects (i.e. a different number of occurrences between total and partial niches due to the nested nature of the data), we randomly resampled the same number of occurrence data for the total niche as we had available for a given partial niche. We repeated the analysis for 100 subsamples of the total niche, thus obtaining a distribution of 100 Schoener's *D*-values and *P*-values for the niche similarity test for each time frame.

Species distribution models

We considered each time frame separately to estimate a climate envelope (hereafter partial-SDM), and then pooled all data together to obtain a total climate envelope (hereafter total-SDM) that should provide a closer approximation of each species' fundamental niche (Nogués-Bravo, 2009) when compared to the partial-SDMs. For each time frame, we generated a set of 2000 random points that were used as pseudo-absences together with observed presences to calibrate the SDMs. We excluded the time frames for 12,000 and 13,000 yr BP for beech because only three observations of presence were available in the former and none in the latter (Table 1).

To model species distributions we used an ensemble forecasting approach as implemented in BIOMOD (Thuiller *et al.*, 2009). We considered the following seven models: (1) generalized linear models (GLMs); (2) generalized additive models (GAMs); (3) artificial neural networks (ANNs); (4) generalized boosted models (GBMs); (5) random forests (RFs); (6) flexible discriminant analysis (FDA); and (7) multivariate adaptive regression spline (MARS). The main references for each model are available in Thuiller *et al.* (2009).

We calibrated models using a random sample of 80% of the original data, and model performance was assessed using the remaining 20% of the data (internal evaluation) by measuring the area under the curve (AUC) of the receiver operating characteristic (ROC) curve (Swets, 1988). The entire procedure was repeated 10 times allowing for the calculation of an average AUC value for each model type and for each time frame. A final model (one for each modelling technique considered) was calibrated using the entire database for each time frame (Thuiller *et al.*, 2009). SDMs with internal AUC value greater than 0.7 (Swets, 1988) were projected onto the current and future climate. Then, for each species and for each set of points considered, we obtained a final consensus forecast of species distribution calculated as the weighted average of all available models (weights for each model based on the respective AUC score as in Marmion *et al.*, 2009). The consensus forecast projected onto the current climate was evaluated by fitting the AUC values against modern pollen data (external evaluation). We transformed the continuous consensus forecasts for both current and future climate into binary predictions of species presence and absence using a threshold that maximizes the sum of sensitivity and specificity as measured against the modern pollen data (Nenzén & Araújo, 2011).

RESULTS

Niche expansion through time

The first two axes of the PCA calibrated over all time frames and considering all variables accounted for 88% of the total variability in climate (51% of variability along the first axis, 37% along the second axis; Fig. 1d), with two main gradients, one corresponding to precipitation and one to temperature. All time frames provided a contribution to the total niche, which increases in size whenever a new time frame is considered (Fig. 1a–c). For silver fir, the oldest occurrences (Fig. 1a) provided an almost complete coverage of the temperature gradient exploited by the species. In contrast, the response along the precipitation gradient during the earliest available time frames was characterized by presences at both the drier and wetter ends of the gradient, whereas presence at intermediate precipitation values only occurred later in time. Beech (Fig. 1b) showed an almost continuous expansion through time along both the precipitation and the temperature gradients, occupying areas with lower precipitation and intermediate temperature in the earlier time frames and expanding towards wetter and both colder and hotter conditions in the time frames closer to the present. Spruce (Fig. 1c) had the widest niche and occurred at the lowest temperature values, while overall it showed a pattern comparable to the one outlined for silver fir.

The overlap between the partial and the total niche increased through time for all species (Fig. 2), being low at 13,000 yr BP and reaching a plateau towards 5000–7000 yr BP (Fig. 2). For beech and fir the oldest time frames showed a low average Schoener's *D* (0.01 and 0.12, respectively), while the overlap was higher for spruce (average Schoener's $D = 0.48$). Maximum overlaps were reached between 2000 and 3000 yr BP (average Schoener's $D = 0.86$ at 2000 yr BP for silver fir; average Schoener's $D = 0.80$ at 3000 yr BP for beech; average Schoener's $D = 0.80$ at 3000 yr BP for spruce). Spruce showed the highest fluctuations in the overlaps between partial and total niches (Fig. 2). Niche similarity was statistically significant for all comparisons, with the exception of three instances for beech (the partial niches between 11,000 and 13,000 yr BP; average *P*-values 0.07, 0.06 and 0.11, respectively), two for fir (11,000 and 13,000 yr BP; average *P*-values 0.06 and 0.11, respectively) and one for spruce (13,000 yr BP; average *P*-value 0.05).

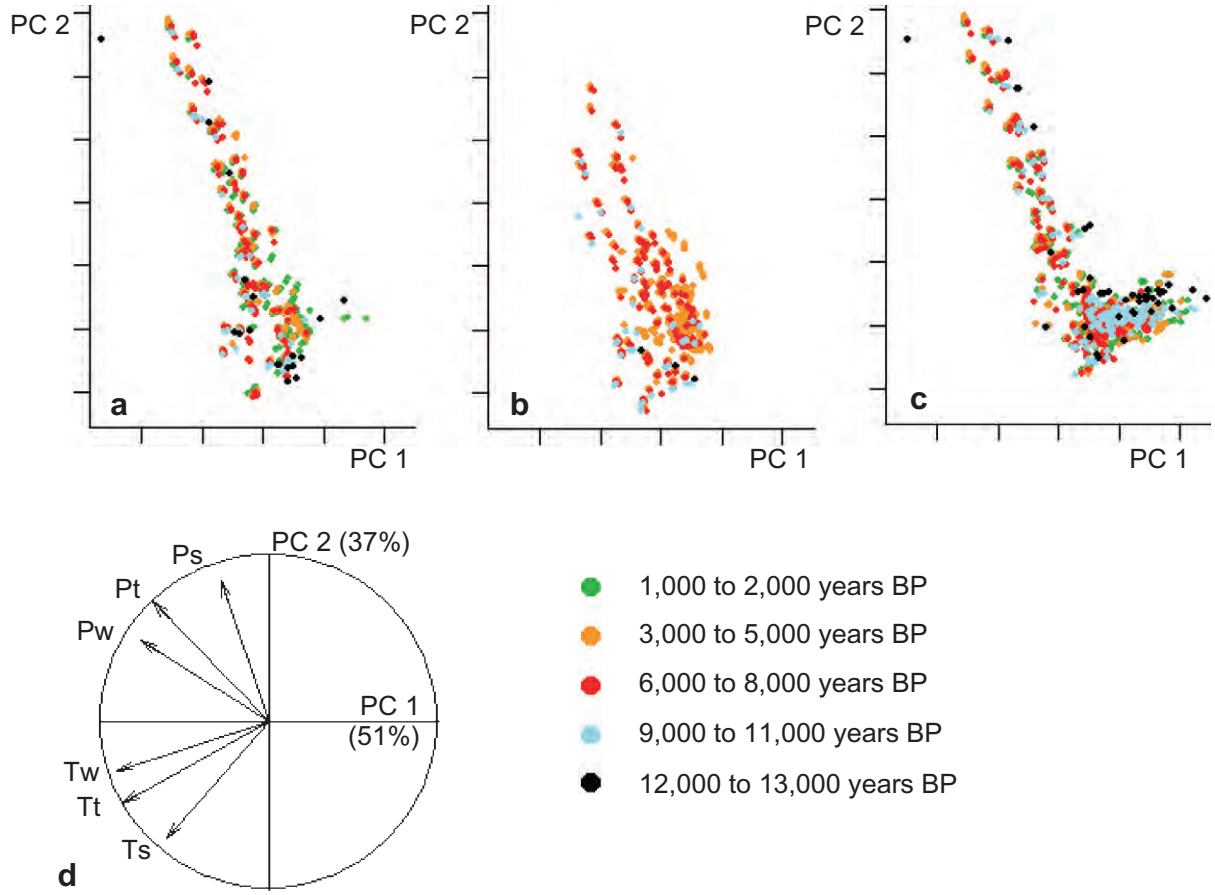


Figure 1 Species niches (a, silver fir; b, beech; c, spruce) projected in the environmental space defined by a principal components analysis calibrated on all environmental variables and all time frames (see text for more details). PC 1 (principal component 1) and PC 2 (principal component 2) represent the first two principal components, explaining 88% of the total variability. Dots in a, b and c represent species distribution in the environmental space with different colours for different time frames. (d) Correlation circle with the variance contribution of axis 1 and 2. Ps = summer precipitation; Pt = annual precipitation; Pw = winter precipitation; Ts = summer mean temperature; Tt = annual mean temperature; Tw = winter mean temperature.

Species distribution models

We generated 77 SDMs for each species and for each time frame (7 modelling techniques by 10 repetitions plus 7 models calibrated on the entire database), for a total of 1001 models for silver fir and spruce, and 693 models for beech. On average, internal evaluation provided good results (*sensu* Swets, 1988) with most AUC values being > 0.85 (Table S1). No modeling algorithm was consistently worse than any other, with FDA having on average the lowest AUC value for silver fir, MARS having the lowest value for beech and ANN for spruce (Table S1). A GAM provided on average the best fit for both silver fir and beech, and was ranked as the third highest algorithm for spruce (GBM being the best model for the latter).

The AUC values calculated by comparing modern pollen data against the consensus forecasts of species distribution under current climate revealed that all models can be considered 'useful' (*sensu* Swets, 1988) with the exception of the partial-SDM obtained for silver fir and spruce calibrated from fossils from 13,000 yr BP (Table 2). In general, model accuracy increased with time frames closer to the present and with increasing sample sizes. Maximum AUC values were obtained by calibrating data from 2000 yr BP for silver fir, from 3000 yr BP for beech, and from 4000 yr BP for spruce. The total-SDM was the best model in the external evaluation for spruce only (Table 2).

For fir and beech, total annual precipitation and summer precipitation were by far the most important variables across all time frames (Tables 3 & S2), followed by summer mean temperature for fir and by annual mean temperature for beech. For spruce, summer precipitation was consistently the most important variable, followed by summer mean temperature and winter mean temperature (Tables 3 & S2).

The partial-SDMs gave different results when projected over the current climate (Fig. S2), suggesting change in the species realized niches through time (Fig. 2). For each of the three species considered, the total-SDMs (Fig. 3a, c, and e) predicted a climatic suitability larger than any of the partial-SDMs (Fig. 3, Table 4), with the only exception of the partial-SDM for the silver fir calibrated with pollen data from 12,000 yr BP. This was

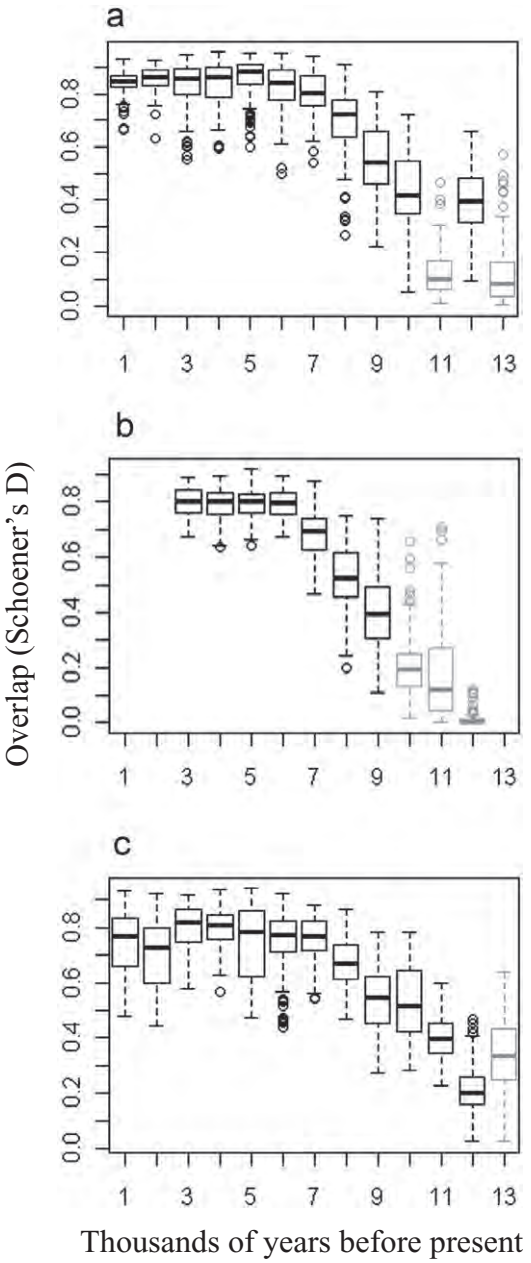


Figure 2 Comparison between partial niches and total niche for all species (a, silver fir; b, beech; c, spruce). Box plots represent Schoener's *D*-values measured considering 100 random subsamples of the total niche with a sample size equivalent to the occurrence data available for each time frame. Light gray boxes show non-significant similarities.

Table 2 Area under the curve (AUC) values calculated for each model using modern pollen data (external evaluation).

Years before present	*Abies alba*	*Fagus sylvatica*	*Picea abies*
13,000	0.664	–	0.625
12,000	0.843	–	0.904
11,000	0.844	0.780	0.867
10,000	0.735	0.739	0.742
9000	0.767	0.807	0.853
8000	0.843	0.842	0.817
7000	0.834	0.879	0.830
6000	0.830	0.922	0.855
5000	0.819	0.809	0.867
4000	0.810	0.957	0.906
3000	0.804	0.959	0.889
2000	0.850	–	0.882
1000	0.829	–	0.823
All years	0.700	0.897	0.917

Table 3 Relative importance of climate variables for silver fir, beech and spruce. Only the most important variable is shown (complete results in Table S1). Results for unreliable species distribution models (as measured through area under the curve values; see the text for more details) are not reported.

Years before present	*Abies alba*	*Fagus sylvatica*	*Picea abies*
13,000	–	–	–
12,000	PRCS	–	PRCS
11,000	PRCT	TAVT	PRCS
10,000	PRCT	TAVT	PRCS
9000	PRCT	PRCT	PRCS
8000	PRCT	PRCS	PRCS
7000	PRCT	PRCS	PRCS
6000	PRCT	PRCS	PRCS
5000	PRCT	PRCS	PRCS
4000	PRCT	PRCT	PRCS
3000	PRCS	PRCT	PRCS
2000	PRCT	–	PRCS
1000	PRCS	–	PRCS
All years	PRCT	PRCS	PRCS

TAVW, winter average temperature; TAVT, average annual temperature; PRCT, total annual precipitation; PRCS, summer precipitation.

especially the case at higher latitudes (Fig. S3). In general, the areas inside the current distribution ranges were characterized by higher suitability values in the total-SDM than in any partial-SDM (Table 4). Moreover, the total-SDMs for the three species also predicted high climate suitability outside of the current distribution range (especially for silver fir and beech), suggesting that: (1) the species considered might not be currently in equilibrium with climate or that (2) the balance of competition has changed.

No matter which climate scenario is considered, the future climate suitability as obtained from the total-SDM for the three species (Fig. 4) is projected to decrease substantially, while a clear northward shift in the species distribution is predicted. The highest decrease is projected for spruce, which will lose suitability in almost 58% of the study area under the A2 scenario (compared with the current potential distribution obtained with the same total-SDM). The lowest decrease is projected for fir, which under the B1 scenario will lose suitability in almost 40% of the study area. When projected into the future, the

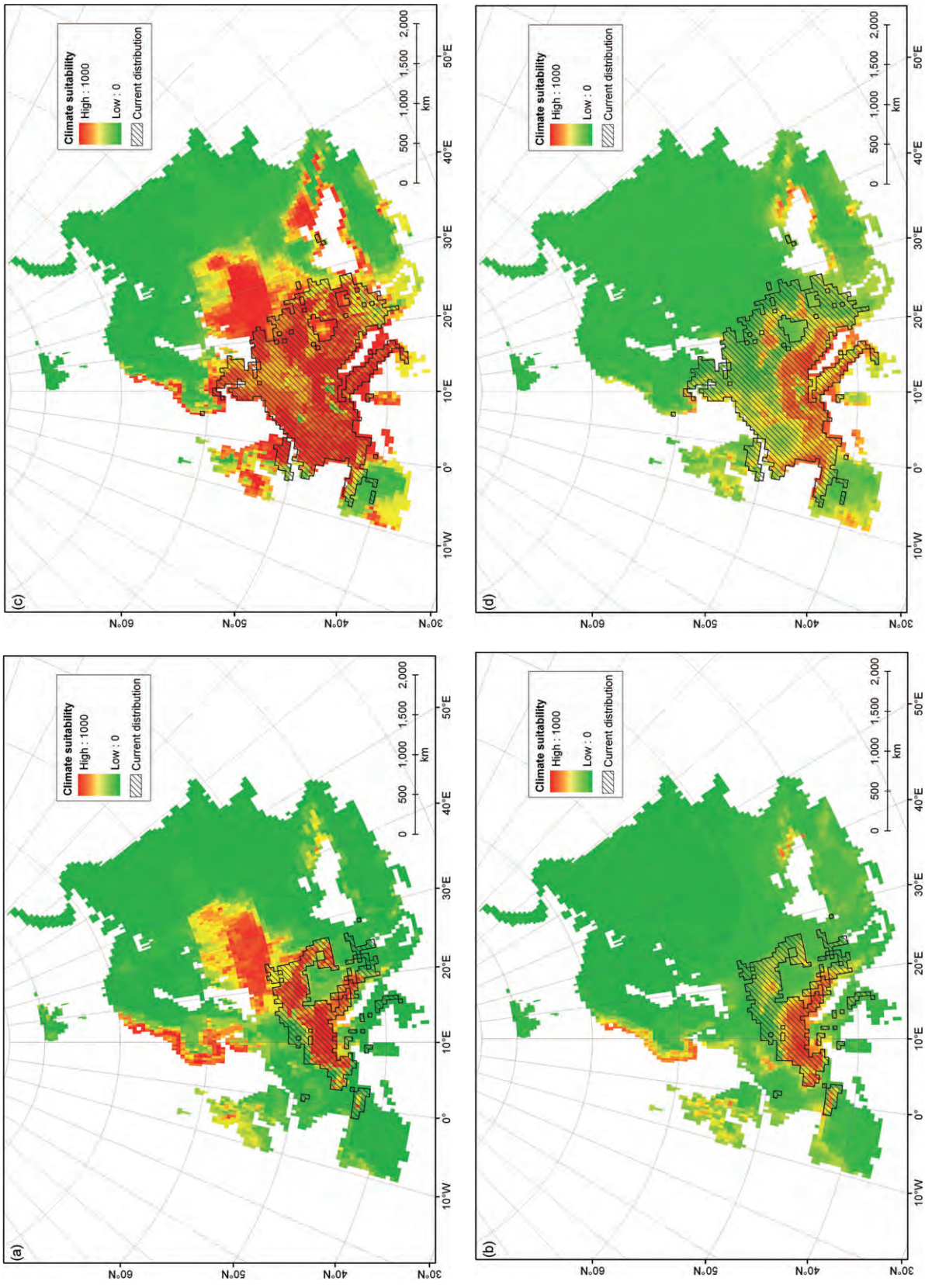


Figure 3 Total species distribution models (total-SDMs) projected on current climate for silver fir (a), beech (c) and spruce (e) and average climate suitability calculated considering all partial-SDMs projected on the current climate (b, silver fir; d, beech; f, spruce). Projection: Albers equal area conic. Datum: European 1950.

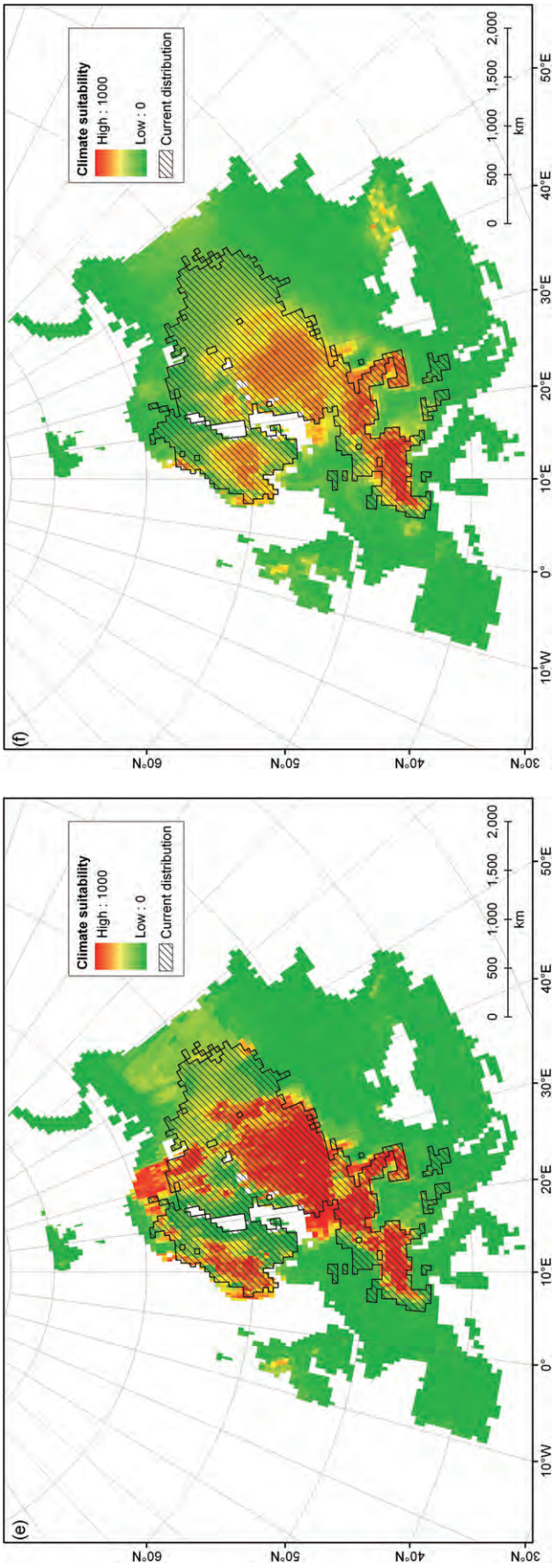


Figure 3 Continued

Table 4 Percentage of the total European land area (% Europe) and of the actual distribution range (% Range) predicted as presences. Presence was obtained from modelled probabilities by applying a threshold that maximizes the sum of sensitivity and specificity as measured against modern pollen data. Results for unreliable species distribution models (as measured through area under the curve values; see the text for more details) are not reported.

	Abies alba		*Fagus sylvatica*		*Picea abies*	
Years before present	% Europe	% Range	% Europe	% Range	% Europe	% Range
13,000	–	–	–	–	–	–
12,000	16.5	45.5	–	–	30.8	65.2
11,000	7.1	29.6	8.7	17.1	29.8	64.0
10,000	3.0	17.8	15.7	27.4	11.1	27.8
9000	6.4	26.7	13.6	26.6	12.3	34.1
8000	8.4	39.1	12.6	30.5	14.8	38.6
7000	8.1	40.5	8.4	27.7	15.7	41.6
6000	6.9	39.8	13.0	42.7	16.0	42.3
5000	9.3	45.5	12.6	41.8	20.2	53.6
4000	8.1	52.9	15.9	63.8	21.1	56.5
3000	6.7	46.9	17.2	66.3	18.5	49.6
2000	9.8	66.4	–	–	19.4	49.6
1000	9.1	60.7	–	–	16.9	44.5
All years	15.5	61.1	37.3	91.7	34.8	75.0

partial-SDMs for the three species provided potential distributions that were almost always underestimates compared with those obtained from the total-SDM (especially for beech; Fig. 4); again results were highly variable from one time frame to another (Figs S4 & S5).

DISCUSSION

Our results demonstrate that the realized climatic niche of the species considered in this study may have changed sufficiently through the Holocene that, when used in bioclimatic SDMs to project current and future distributions, different results are obtained depending upon which observed niche(s) are used to calibrate the models. This may be due to several factors, including changes in competition through time, or non-analogue climates that occurred in the past. Of particular interest is the pattern of change in bioclimatic niches that we measured. Beech had an almost continuous expansion through time along both the precipitation and the temperature gradients. This may reflect that known glacial refugia for beech located in the Pyrenees, the western Alps, the eastern Alps, southern Italy, the Balkans and the Carpathians (Magri *et al.*, 2006) represent restricted climate space, compared with the potential climate space beech can occupy. Its Holocene expansion was accompanied by cooler and wetter conditions, occurring around 8000 yr BP in southern Central Europe and between 4000 and 2000 yr BP in northern Europe (Tinner & Lotter, 2006). For both fir and spruce, the oldest occurrences provided an almost complete coverage of the temperature gradient exploited by the species as measured by the total niche, but not of the precipitation gradient; again this probably reflects the location of their glacial refugia (Tollefsrud *et al.*, 2008; Liepelt *et al.*, 2009). The expansion of the two species was accompanied by changing climatic factors, like wetter conditions around 8000 yr BP (Tinner & Lotter, 2006), but was probably also influenced by competition with beech and other species and by fire and other human disturbances (Lang, 1994).

In most cases, the partial climatic niches (i.e. the climatic niche estimated using one time frame at a time) showed a significant overlap with the total niche (i.e. the climatic niche estimated using all data at once), with maximum values of average Schoener's *D* being equal or greater than 0.8 (i.e. 80% of overlap) for all species, and with nearly all values being above 0.5 after 10,000 yr BP for all species (Fig. 2). Yet, we never obtained an average Schoener's *D* equal to 1, indicating that no individual time frame provided a complete agreement with the total niche. Moreover, when considering only the oldest time frames (i.e. before 10,000 yr BP) no statistically significant Schoener's *D*-values were achieved, indicating that for these time frames and species the partial niches were not more similar to the total niche than random niches built in the same study area. However, the older time frames are also those with lower sample sizes, thus downgrading the power in statistical tests. All partial-SDMs when projected over the current climate gave different patterns of a species' potential distribution (Figs 3 & S2) demonstrating that the effectiveness of bioclimatic SDMs in projecting species distribution varies among species and also among time frames.

Our results parallel, at least in part, those of SDM transferability in geographical space (Randin *et al.*, 2006) and time (Dobrowski *et al.*, 2011), in which the accuracy of SDMs varies as a function of several factors, including species ecology, geographical area and time frame considered. A number of factors can lead to changes in the realized niche of a species, including dispersal limitations, competition and other historical and biological factors such as human disturbance (Meier *et al.*, 2011a; Normand *et al.*, 2011). Currently the evidence for niche stability (and conservatism) in the literature is mixed (Wiens & Graham, 2005; Peterson, 2011) and often lacks a distinction between fundamental and realized niche. Starting from a literature review, Peterson (2011) proposed a framework in which the main component in determining niche stability and conservatism is time. Peterson (2011) refers particularly to the evolutionary conservatism of coarse-resolution Grinnellian ecological niches and suggests that recent and short-term events, from decadal and

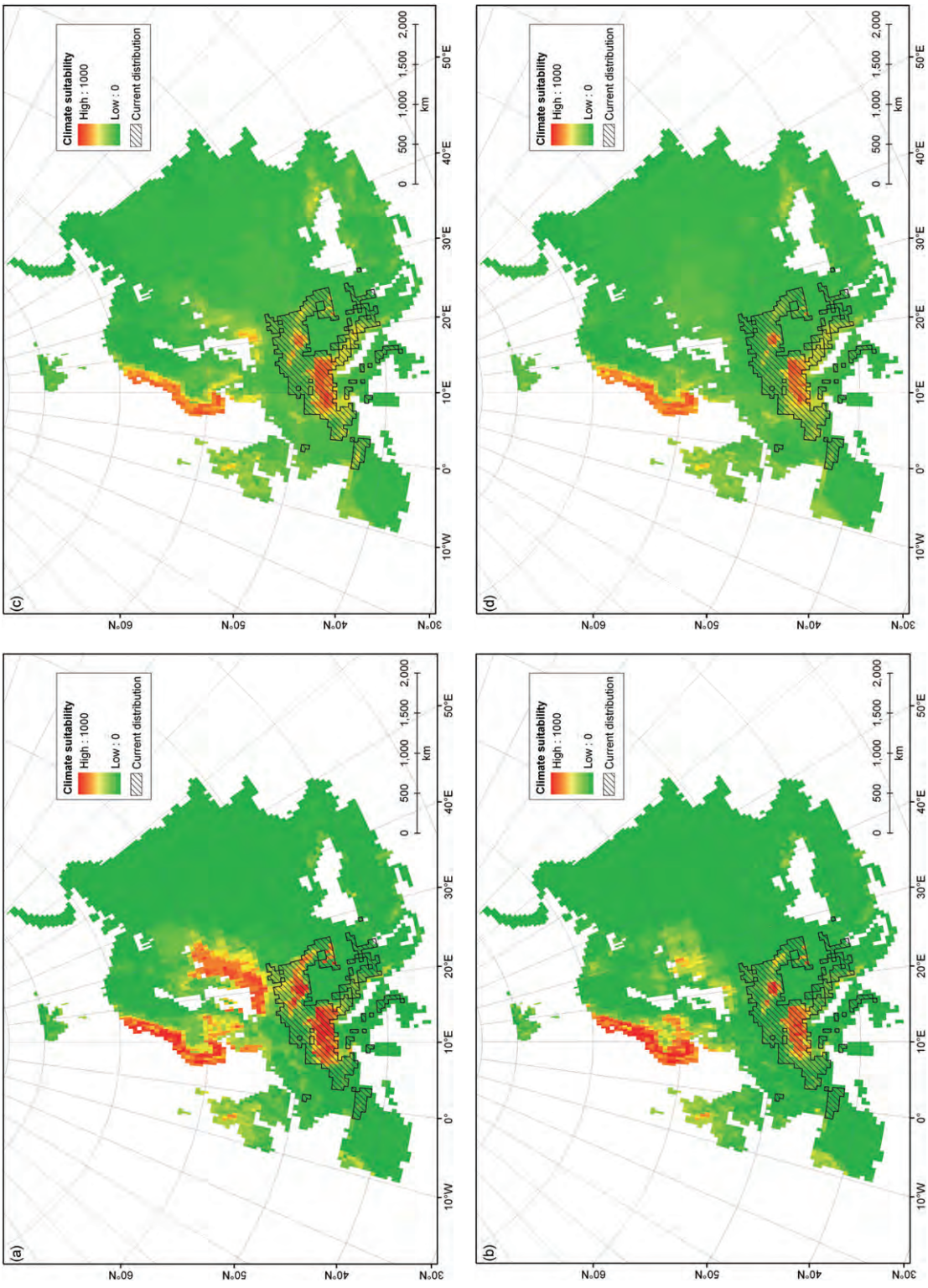


Figure 4 Projected future climate suitability obtained from the total species distribution models (total-SDMs) and average future climate suitability obtained from the partial-SDMs. Detailed results for all partial-SDMs are available in the Supporting Information; averages are given here only to provide a quick illustration of the results: (a) scenario B1, silver fir, total-SDM; (b) scenario A2, silver fir, total-SDM; (c) scenario B1, silver fir, average over all partial-SDMs; (d) scenario A2, silver fir, average over all partial-SDMs; (e) scenario B1, beech, total-SDM; (f) scenario A2, beech, total-SDM; (g) scenario B1, beech, average over all partial-SDMs; (h) scenario A2, beech, average over all partial-SDMs; (i) scenario B1, spruce, total-SDM; (j) scenario A2, spruce, total-SDM; (k) scenario B1, spruce, average over all partial-SDMs; (l) scenario A2, spruce, average over all partial-SDMs. Projection: Albers equal area conic. Datum: European 1950.

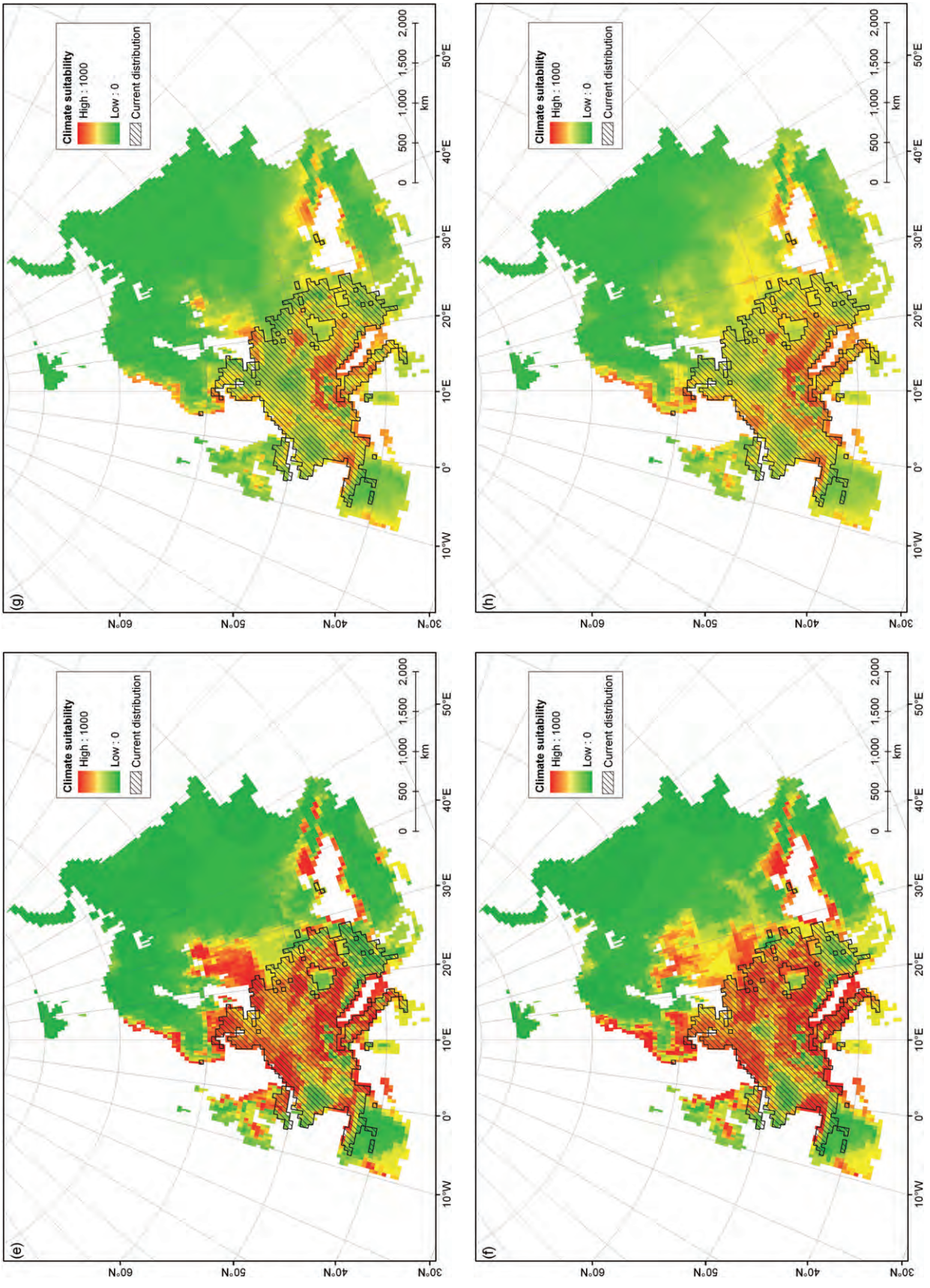


Figure 4 Continued

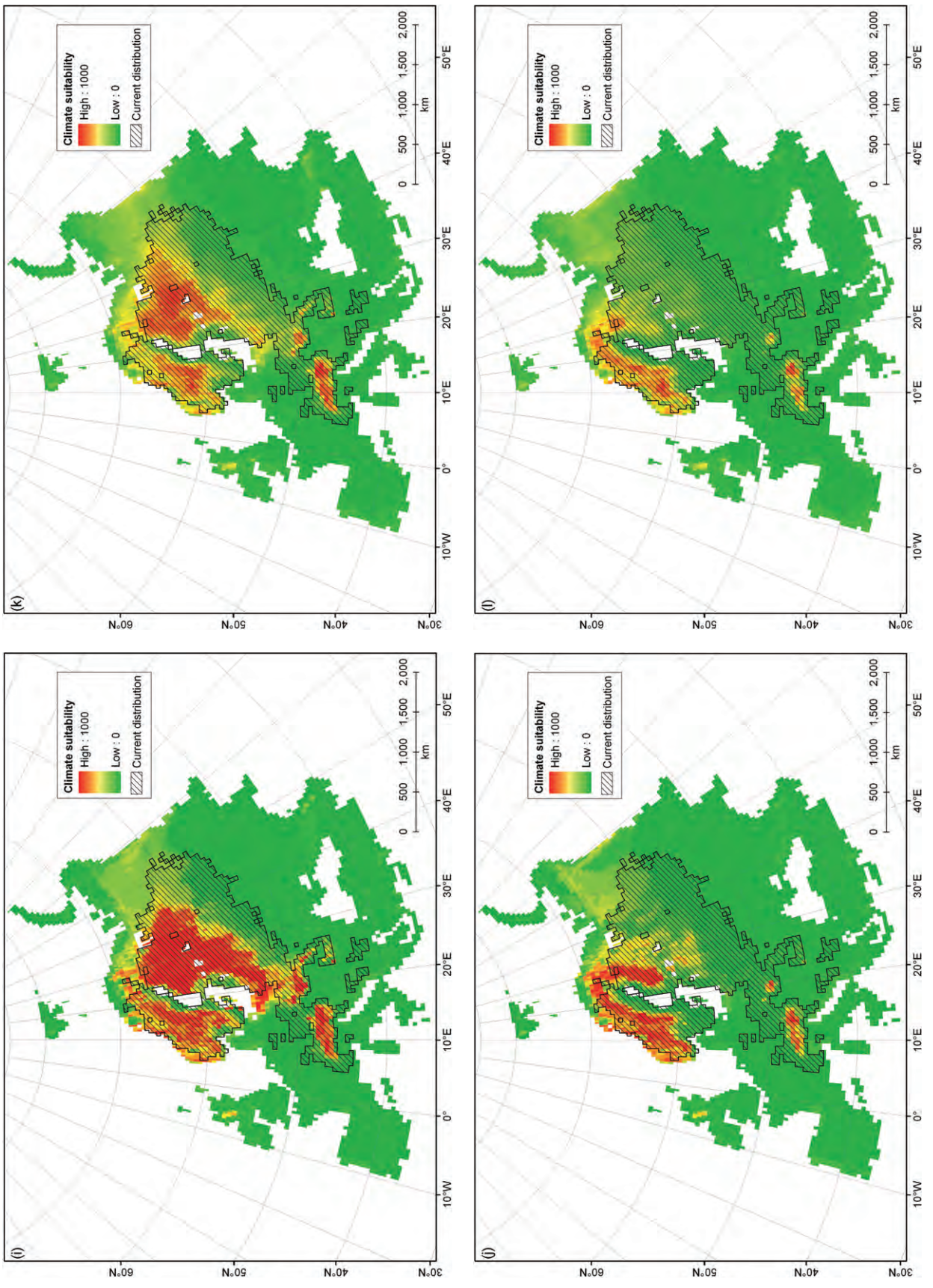


Figure 4 Continued

century-scale species invasions to distributional shifts at the end of the Pleistocene and throughout the Holocene, show a considerable tendency towards niche stability, while longer-term events show increasing degrees of breakdown of stability.

We measured non-significant niche differences through the entire Holocene. However, even if differences in the realized niche were small and non-significant for time frames closer to the present, the effects on potential distribution were important, cautioning against the use of bioclimatic SDMs in predicting species responses to global changes (Meier *et al.*, 2011a,b). Most partial-SDMs predicted high climate suitability values for areas inside the current distribution range (Table 4), suggesting that bioclimatic reconstructions from fossil data are able to fit current species ranges. However, we measured important differences outside the current range, specifically in areas the species are projected to colonize following climate change. This suggests that limited (and not statistically significant) changes in the realized niche may have profound consequences on the projected new distribution of the species.

All SDMs provided a reasonable reconstruction of species niche when compared with data from the literature. In particular, precipitation (total or summer) was consistently the most important variable through all time frames for the three species, a pattern that is confirmed by different studies (e.g. Lang, 1994). Summer and winter mean temperatures were also fairly important for spruce, confirming again what has been measured in the field (Lang, 1994). These climatic controls can have important consequences for projections into future climates, particularly because precipitation patterns are projected to change as much as temperatures, but with far greater uncertainties in amount, direction and spatial patterns (IPCC 2007). This suggests that ensemble approaches are crucial when forecasting under such uncertainties (Araújo & New, 2007).

The partial-SDMs projected to current climate almost always gave high AUC values. For all species (but particularly for spruce), the SDM providing the highest AUC values when projected to current climate and measured against current pollen data was not the one calibrated closest to present, indicating that factors other than climate (e.g. human activities) have been playing an important role in shaping species ranges, at least in the last two or three millennia. However, these results should be considered with caution, given that the temporal trends found in AUC values (Table 2) may be confounded with sample size (Table 1), suggesting that there is at least a possibility that the patterns observed are also influenced by the increased sample size, as opposed to resulting from ecological phenomena pertaining to the species niches.

For all species, the potential current distribution projected from the total-SDM approximated fairly well the southern boundaries of the current distribution, but high suitability values were predicted at the northern and eastern edges, where species are currently absent, probably because of dispersal limitation, increased competition, or because of the interaction of competition and climate. The total-SDM, with the exception of spruce, did not provide the best model when projected under current climate and evaluated against modern pollen data. However, the aim of the total-SDM approach is to provide an estimate of the niche that is closer to the fundamental niche for each species, and thus, by definition, it is possible to obtain an over-estimation of the current species distribution (see above; Colwell & Rangel, 2009).

The projections for future climate show that partial-SDMs systematically produce smaller estimates of the potential species distribution compared with total-SDMs, with a dramatic difference in the case of beech. These results have important consequences for studying the responses of biodiversity to global change. First, we obtained a clear indication that previous results may have underestimated the capacity of tree species to respond to climate change, suggesting that the severity of future threats to biodiversity may have been overestimated in some cases. However, although our projections suggest a potential increase in habitat suitability at higher latitudes, the natural colonization of areas that become suitable under climate change can be slow and difficult. Beech exemplifies a species which has not finished colonizing suitable habitats after the retreat of the ice at the end of the last ice age (Bradshaw & Lindbladh, 2005), and which can be expected to adjust its range only very slowly (Meier *et al.*, 2011b).

The multi-temporal approach we followed has been proposed as an improvement over classical SDMs to project species distributions in time (Nogués-Bravo, 2009; Varela *et al.*, 2009). Compared with previous studies (e.g. Nogués-Bravo *et al.*, 2008; Pearman *et al.*, 2008b) we considered a greater number of time frames, which should have provided a more complete coverage of the species niche (thus more closely approaching the fundamental niche for the entire Holocene). Furthermore, we were able to demonstrate fluctuations of the realized niche in time. Obviously, we did not provide a complete representation of the fundamental niche of any species. In fact, the range of environmental conditions that characterize the observed geographical distribution of a species may fail to reveal the full extent of its fundamental niche, for three main reasons (Colwell & Rangel, 2009): (1) part of a species' fundamental niche does not correspond to any combination of environmental variables; (2) dispersal limitations or (3) species interactions may prevent the occupation of given patches. The total niche estimated through the multi-temporal approach does not provide a complete solution to these three points but at least provides a larger sample of environmental conditions and, presumably, species interactions, and it allows for a relaxation of dispersal limitations, ensuring a better representation of the species ecology overall.

We adopted an ensemble forecasting approach (Thuiller *et al.*, 2009), overcoming some of the shortcomings of previous single modelling approaches (Hijmans & Graham, 2006). A number of limitations still remain. In particular, we considered only a single global circulation model (GCM) for estimating past climate, although results of SDMs can vary substantially with different GCMs (Buisson *et al.*, 2010). It would be interesting to adopt a full ensemble forecasting approach (*sensu* Araújo & New, 2007) considering not only many different statistical modelling algorithms for niche calibration, but also a full suite of GCMs for reconstructing past climate. In any case, even con-

sidering the ideal setting with regard to data availability (for both climate and species), species ecology and modelling framework, the multi-temporal approach, as we outlined here, cannot deal properly with novel climates (Williams & Jackson, 2007). Moreover, with the multi-temporal approach, especially when considering longer time frames, there is also the risk of including different ecotypes (Benito Garzón *et al*., 2011) in the same model, or even changes in the fundamental niche resulting through rapid evolution (Peterson, 2011).

Ideally, in order to predict species responses to climate changes, one should use the physiological limits to obtain a complete representation of the fundamental niche of a species and then constrain it with biotic interaction and dispersal limitation effects (Meier *et al*., 2011a). Hybrid SDMs representing seed dispersal and species interactions are currently available (Engler & Guisan, 2009; Meier *et al*., 2011b), and important research efforts are ongoing in order to understand the best way in which they can be combined with classical static SDMs (Guisan & Rahbek, 2011). However, a clear and complete knowledge of a species' physiology is limited to a few, very particular cases (Kearney & Porter, 2009), and thus a reconstruction of the fundamental niche is currently out of reach for the great majority of species. Building the niche through time can provide a way towards a better understanding of species distribution and ecology in a changing climate.

REFERENCES

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.

Araújo, M.B. & Pearson, R.G. (2005) Equilibrium of species' distribution with climate. *Ecography*, **28**, 693–695.

Benito Garzón, M., Alía, R., Robson, T.M. & Zavala, M.A. (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766–778.

Bradshaw, R.H.W. & Lindbladh, M. (2005) Regional spread and stand-scale establishment of *Fagus sylvatica* and *Picea abies* in Scandinavia. *Ecology*, **86**, 1679–1686.

Broennimann, O. & Guisan, A. (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, **4**, 585–589.

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.-J., Randin, C., Zimmermann, N.E., Graham, C.H. & Guisan, A. (2012) Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, **21**, 481–497.

Buisson, L., Thuiller, W., Casajus, N., Lek, S. & Grenouillet, G. (2010) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, **16**, 1145–1157.

Colwell, R.K. & Rangel, T.F. (2009) Hutchinson's duality: the once and future niche. *Proceedings of the National Academy of Sciences USA*, **106**, 19651–19658.

Dobrowski, S.Z., Thorne, J.H., Greenberg, J.A., Safford, H.D., Mynsberge, A.R., Crimmins, S.M. & Swanson, A.K. (2011) Modeling plant ranges over 75 years of climate change in California, USA: temporal transferability and species traits. *Ecological Monographs*, **81**, 241–257.

Dormann, C.F. (2007) Promising the future? Global change projections of species distributions. *Basic and Applied Ecology*, **8**, 387–397.

Eberling, S., Welk, E., Auge, H. & Bruelheide, H. (2008) Predicting the spread of an invasive plant: combining experiments and ecological niche models. *Ecography*, **31**, 709–719.

Engler, R. & Guisan, A. (2009) MIGCLIM: predicting plant distribution and dispersal in a changing climate. *Diversity and Distributions*, **15**, 590–601.

Guiot, J., Cheddadi, R., Prentice, I.C. & Jolly, D. (1996) A method of biome and land surface mapping from pollen data: application to Europe 6000 years ago. *Palaeoclimates*, **1**, 311–324.

Guisan, A. & Rahbek, C. (2011) SESAM – a new framework for predicting spatio-temporal patterns of species assemblages: integrating macroecological and species distribution models. *Journal of Biogeography*, **38**, 1433–1444.

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

Hijmans, R.J. & Graham, C.H. (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272–2281.

Hutchinson, G.E. (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.

IPCC (Intergovernmental Panel on Climate Change) (2007) *Climate change 2007: the physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the IPCC*. Cambridge University Press, Cambridge, UK.

Jalas, J. & Suominen, J. (1972–1999) *Atlas Florae Europaeae: distribution of vascular plants in Europe*. The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo, Helsinki.

Kearney, M. & Porter, W. (2009) Mechanistic niche modeling: combining physiological and spatial data to predict species ranges. *Ecology Letters*, **12**, 334–350.

van der Knaap, W.O., van Leeuwen, J.F.N., Finsinger, W., Gobet, E., Pini, R., Schweizer, A., Valsecchi, V. & Ammann, B. (2005) Migration and population expansion of *Abies, Fagus, Picea*, and *Quercus* since 15000 years in and across the Alps, based on pollen-percentage threshold values. *Quaternary Science Reviews*, **24**, 645–680.

Lang, G. (1994) *Quartäre Vegetationsgeschichte Europas*. Gustav Fischer, Jena.

Latalowa, M. & van der Knaap, W.O. (2006) Late Quaternary expansion of Norway spruce *Picea abies* (L.) Karst. in Europe according to pollen data. *Quaternary Science Reviews*, **25**, 2780–2805.

Liepelt, S., Cheddadi, R., de Beaulieu, J.L., Fady, B., Gomory, D., Hussendorfer, E., Konnert, M., Litt, T., Longauer, R.,

Terhürne-Berson, R. & Ziegenhagen, B. (2009) Post-glacial range expansion and its genetic imprints in *Abies alba* (mill.): a synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology*, **153**, 139–149.

Magri, D., Vendramin, G.G., Comps, B., Dupanloup, I., Geburek, T., Gömöry, D., Latalowa, M., Litt, T., Paule, L., Roure, J.M., Tantau, I., van der Knaap, W.O., Petit, R.J. & de Beaulieu, J.L. (2006) A new scenario for the Quaternary history of European beech populations: palaeobotanical evidence and genetic consequences. *New Phytologist*, **171**, 199–221.

Maiorano, L., Falcucci, A., Zimmermann, N.E., Psomas, A., Pottier, J., Baisero, D., Rondinini, C., Guisan, A. & Boitani, L. (2011) The future of terrestrial mammals in the Mediterranean basin under climate change. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2681–2692.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modeling. *Diversity and Distributions*, **15**, 59–69.

Meier, E.S., Edwards, T.C., Kienast, F., Dobbertin, M. & Zimmermann, N.E. (2011a) Competition patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica*. *Journal of Biogeography*, **38**, 371–382.

Meier, E.S., Lischke, H., Schmatz, D.R. & Zimmermann, N.E. (2011b) Climate, competition and connectivity affect future migration and ranges of European trees. *Global Ecology and Biogeography*, **21**, 164–178.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. & New, M. (2004) A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100). *Tyndall Centre Working Paper*, **55**, 1–25.

Nenzén, H.K. & Araújo, M.B. (2011) Choice of threshold alters projections of species range shifts under climate change. *Ecological Modelling*, **222**, 3346–3354.

Nogués-Bravo, D. (2009) Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, **18**, 521–531.

Nogués-Bravo, D., Rodríguez, J., Hortal, J., Batra, P. & Araújo, M. (2008) Climate change, humans, and the extinction of the woolly mammoth. *PLoS Biology*, **6**, e79.

Normand, S., Ricklefs, R.E., Skov, F., Bladt, J., Tackenberg, O. & Svenning, J.C. (2011) Postglacial migration supplements climate in determining plant species ranges in Europe. *Proceedings of the Royal Society B: Biological Sciences*, **278**, 3644–3653.

Pearman, P., Guisan, A., Broennimann, O. & Randin, C. (2008a) Niche dynamics in space and time. *Trends in Ecology and Evolution*, **23**, 149–158.

Pearman, P., Randin, C., Broennimann, O., Vittoz, P., van der Knaap, W., Engler, R., Le Lay, G., Zimmermann, N.E. & Guisan, A. (2008b) Prediction of plant species distributions across six millennia. *Ecology Letters*, **11**, 357–369.

Peterson, A.T. (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, **38**, 817–827.

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

Singarayer, J.S. & Valdes, P.J. (2010) High-latitude climate sensitivity to ice-sheet forcing over the last 120 kyr. *Quaternary Science Reviews*, **29**, 43–55.

Svenning, J.-C., Normand, S. & Skov, F. (2008) Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, **31**, 316–326.

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

Terhürne-Berson, R., Litt, T. & Cheddadi, R. (2004) The spread of *Abies* throughout Europe since the last glacial period: combined macrofossil and pollen data. *Vegetation History and Archaeobotany*, **13**, 257–268.

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

Tinner, W. & Lotter, A.F. (2006) Holocene expansions of *Fagus silvatica* and *Abies alba* in central Europe: where are we after eight decades of debate? *Quaternary Science Reviews*, **25**, 526–549.

Tollefsrud, M.M., Kissling, R., Gugerli, F., Johnsen, Ø., Skrøppa, R., Cheddadi, R., van der Knaap, W., Latalowa, M., Terhurne-Berson, R., Litt, T., Geburek, T., Brochmann, C. & Sperisen, C. (2008) Genetic consequences of glacial survival and postglacial colonization in Norway spruce: combined analysis of mitochondrial DNA and fossil pollen. *Molecular Ecology*, **17**, 4134–4150.

Varela, S., Rodríguez, J. & Lobo, J.M. (2009) Is current climatic equilibrium a guarantee for the transferability of distribution model predictions? A case study of the spotted hyena. *Journal of Biogeography*, **36**, 1645–1655.

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

Williams, J. & Jackson, S. (2007) Novel climates, no-analog plant communities, and ecological surprises: past and future. *Frontiers in Ecology and Evolution*, **5**, 475–482.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1 Current species distribution and modern pollen samples available for the three studied tree species.

Figure S2 Species distribution models calibrated with fossil data and projected on current climate for the three studied tree species.

Figure S3 Binary projections of species distribution under current climate for the three studied tree species.

Figure S4 Species distribution models calibrated with fossil data and projected on future for the three studied tree species.

Figure S5 Binary projections of species distribution under future for the three studied tree species.

Table S1 Average area under the curve values calculated in the internal model evaluation for each modelling algorithm considered.

Table S2 Average relative variable importance for the three studied tree species.

BIOSKETCHES

The ECOCHANGE project (http://www.ecochange-project.eu) was supported under the 6th Framework Programme of the European Union (contract number FP6-036866). The project is aimed at assessing the capacity of ecosystems to supply humans with required goods and services and to buffer against climate and land-use change, with a particular focus on the improvement of models and the generation of new climate change projections on European biodiversity.

Author contributions: L.M., P.B.P. and A.G. (co-financed by NSF grant number 31003A-125145 'BIOASSEMBLE') designed the study; R.C., H.L. and J.S.S. collected the data; L.M. conducted the analyses with the help of L.P., B.P. (co-financed by the Swiss National Center for Competence in Research 'Plant Survival' project), N.E.Z. and B.I.H. (co-financed by doctoral contract POSDRU 6/1.5/S/3 - 'Doctoral studies: through science towards society'). L.M. and A.G. wrote the manuscript with the help of all co-authors.

Editor: Solomon Dobrowski