

$$\frac{1}{2} \sum_j (b_j)^2$$

$$\pi(x) =$$

$$\frac{1}{K} \int_{q_0(x)}^{\log} \sum_{i=1}^n \beta_i = 0$$

$$\pi_s [1/s] / \pi_s [1/s] =$$

$$\pi(x) = \prod_{1 \leq l \leq L}$$

$$J_k = \left\{ \left| \frac{dx_j}{dt} \right| > 0 \right\} \cap B_i$$

$$\eta(N, E, T) = G - \eta(E, T)$$

$$P(Y=1, X=q) = \frac{b(q)}{Z}$$

$$\hat{N} = \frac{1}{\alpha} \left[1 - \frac{K}{G} \right]$$

$$RE(\pi \| q) \leq RE(\pi \| q^*)$$

$$\min_{\lambda \in \mathbb{R}} \sum_{i=1}^n \lambda_i \sum_{x \in \mathcal{X}} p_1(x) p_2(x)$$

$$a_j, b_j = [\pi \| q]$$

$$\min_{\lambda \in \mathbb{R}} \sum_{i=1}^n \lambda_i$$

$$P_{\text{opt}} =$$

$$\frac{b(q)}{G} \pi(x) = K$$

$$A \frac{dx_j}{dt}$$

$$A$$

$$\frac{dx_j}{dt}$$

$$P_2(x)$$

$$\frac{1}{2} N \log q$$

$$\sum_j \beta_j$$

$$\sum_{x \in \mathcal{X}} p_1(x) p_2(x)$$

$$RE$$

$$\log Z$$

$$\sum_j \beta_j$$

$$\sum_{i=1}^n \lambda_i$$

$$RE(\pi \| q)$$

$$RE(\pi \| q) \leq F$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\frac{1}{2} \sum_j \beta_j$$

$$\log Z$$

$$\sum_j \beta_j$$

$$\sum_{i=1}^n \lambda_i$$

$$RE(\pi \| q)$$

$$RE(\pi \| q) \leq F$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\frac{1}{2} \sum_j \beta_j$$

$$\log Z$$

$$\sum_j \beta_j$$

$$\sum_{i=1}^n \lambda_i$$

$$RE(\pi \| q)$$

$$RE(\pi \| q) \leq F$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\frac{1}{2} \sum_j \beta_j$$

$$\log Z$$

$$\sum_j \beta_j$$

$$\sum_{i=1}^n \lambda_i$$

$$RE(\pi \| q)$$

$$RE(\pi \| q) \leq F$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

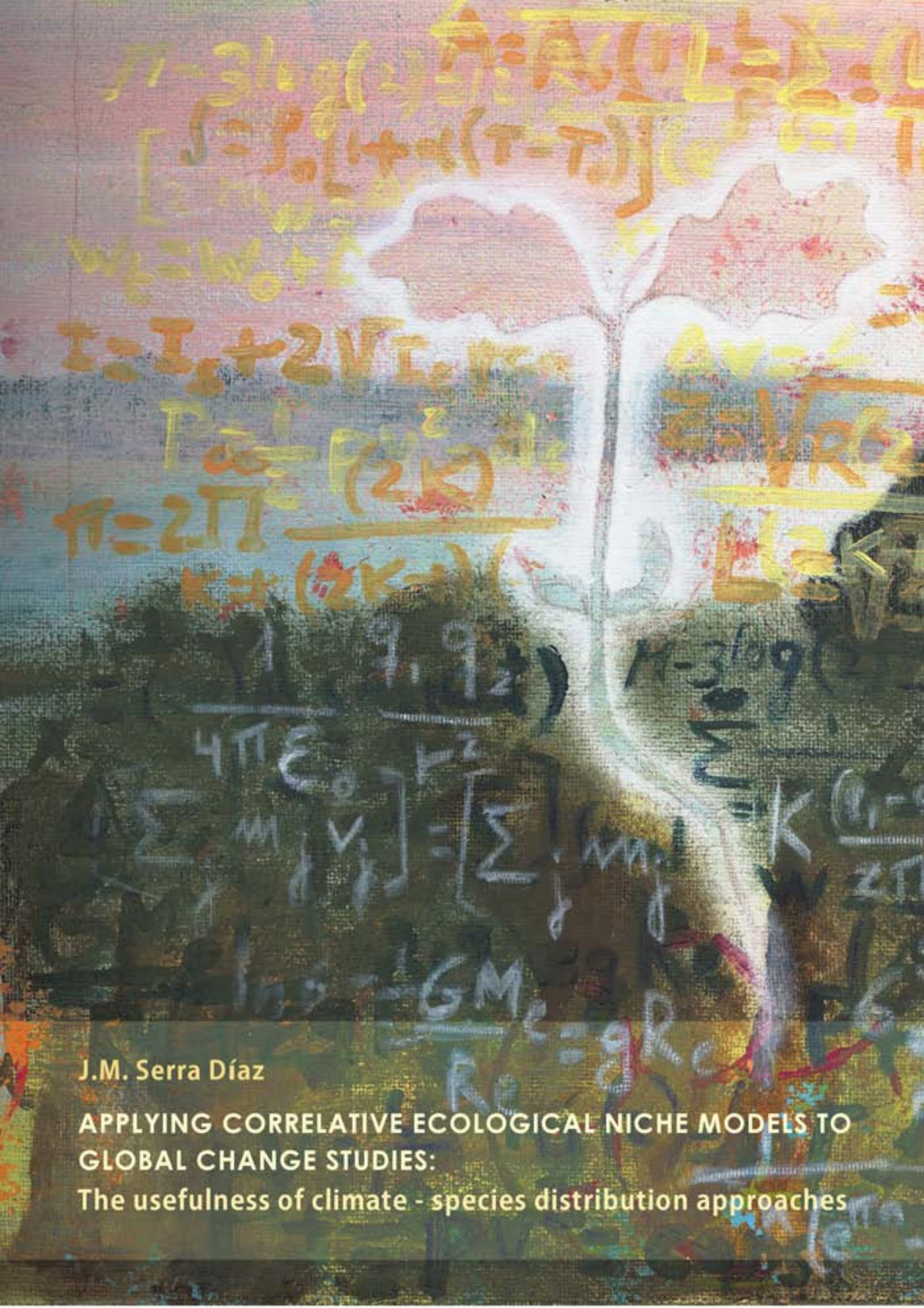
$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$

$$\beta_j = (B/\sqrt{m})$$



J.M. Serra Díaz

APPLYING CORRELATIVE ECOLOGICAL NICHE MODELS TO
GLOBAL CHANGE STUDIES:
The usefulness of climate - species distribution approaches

TABLE OF CONTENTS

1.INTRODUCTION	11
1.1 Preliminaries.....	11
Knowing, understanding and predicting species distributions under global change: A cross-disciplinary challenge.....	11
1.2 Setting the scene of species distribution models: a very brief summary (or not).	15
What drives species distribution? Not so easy choosing predictors.....	16
What is my data like? Choosing or adapting to data approaches and sampling strategies.....	18
Embracing a whole ecosystem of algorithms.....	19
Spatial and temporal inference	22
Not so obvious thresholds.....	23
Is my model an accurate model?.....	24
Assuming a utopia?	26
1.3 Theoretical framework: The history and (the continuous) evolution of the niche theory, and its relationship with species distribution.	27
1.4 Modeling framework.....	31
1.5 Applied framework of CENM in the face of global change	33
Biodiversity monitoring and discovering.....	33
Conservation planning and management	34
Species invasions	35
Epidemiology and disease transmission studies	35
1.6 Aims and scope.....	37
1.7 References	39
2.FIRST STEPS.....	54
2.1 The topoclimatic Atlas of the Iberian Peninsula.....	54
2.2 Material and methods.....	55
Presence-Absence data of species.	55
Selection of topographic and climatic variables.....	55
Evaluation	56

Future scenarios: General Circulation Models (GCM).....	56
2.3 References and software.....	57
2.4 Acknowledgements and participants	59
3.ON CURRENT DISTRIBUTIONS	60
3.1 Introduction.....	60
3.2 Material and methods.....	64
Niche-based models (NBM).....	64
Process-based models (PBM)	66
Study design	67
3.3 Results	69
General patterns of congruence and incongruence between growth and suitability values.....	69
Incongruence of virtual forests and geographical distribution	70
Physiological responses: Mean leaf life (MLL) and leaf area index (LAI).....	72
Differences between Niche-based model techniques and the role of the threshold value.....	74
3.4 Discussion	75
3.5 References.....	80
4. ON THE IMPORTANCE OF (PHYSIOLOGICAL) MECHANISMS IN CLIMATE CHANGE PROJECTIONS	88
4.1 Introduction.....	88
4.2 Material and methods.....	91
Selected region and study species	91
Climatic data.....	92
Niche-based models	93
Ecosystem model GOTILWA+	96
Experimental setup.....	97
4.3 Results	98
4.4 Discussion	104
4.5 References	111
5. ON PROJECTIONS OF MULTIPLE SPECIES: AGGREGATED VERSUS INDIVIDUAL PLANT SPECIES DISTRIBUTIONS.....	124
5.1 Introduction.....	124
5.2 Material and methods.....	127
Selection of variables.....	128
Generalized linear models for <i>Abies alba</i> , <i>Fagus sylvatica</i> and <i>A. alba-F. sylvatica</i> co-occurrence	130

Future climatic scenarios: downscaling global circulation models	130
5.3 Results	131
5.4 Discussion	136
5.5 References	141
6. ON FACING POTENTIAL NICHE CHANGES	149
6.1 Introduction	149
6.2 Material and methods	152
6.3 Results	156
6.4 Discussion	162
6.5 References	165
7. ON CONSIDERING THE TEMPORAL DIMENSION	171
7.1 Introduction	171
7.2 Material and methods	173
7.3 Results	176
7.4 Discussion	183
7.4 References	186
8. FINAL REMARKS AND LESSONS LEARNED	190
8.1 Reflecting on the use of correlative Ecological Niche Models for global change studies	190
8.2 The ongoing and future modeling scene and the role for CENM	199
8.3 So far, what have I learned along the way? Bits and pieces about CENM in a nutshell	203
Do not be afraid of comparing apples and oranges!	203
Fit the model for the purpose and, if possible, do not run out of time.	203
Watch out going beyond your predictions	203
8.4 References	204
AGRAÏMENTS I RECONeixEMENTS	212
APPENDIX	213
APPENDIX 1	214
APPENDIX 1.1. Literature extension	214
APPENDIX 2	238
APPENDIX 2.1. Results from niche-based models. (NBM)	238
APPENDIX 2.2 Geography of mean leaf life (MLL) simulations in <i>Pinus halepensis</i>	241
APPENDIX 2.3. Analysis of variance results between NPP and suitable/unsuitable virtual forests for <i>Quercus ilex</i> and <i>Pinus sylvestris</i>	242

APPENDIX 3.....	243
APPENDIX 3.1 Extrapolation in the environmental space.....	243
APPENDIX 4.....	247
APPENDIX 4.1 Model descriptions.....	247
APPENDIX 4.2 Model Accuracy.....	249
APPENDIX 4.3 Metrics description	252

INDEX OF FIGURES

CHAPTER 1

Figure 1.1 Modeling sequence of correlative ecological niche models.....	16
Figure 1.2 Relationship between algorithm and data approach with the type of niche being modeled. Modified from Jimenez-Valverde et al. (2008).....	22
Figure 1.3 Conceptual framework of species distribution of the BAM diagram	30
Figure 1.4 Model classification between trade-off scheme of 3 properties: Precision, generality and reality. After Levins (1966) and inspired by Guisan and Zimmermann (2000).....	32

CHAPTER 2

Figure 2.1 Web display of the Atlas of Topo-climatic suitability.....	54
-----------------------------------------------------------------------	----

CHAPTER 3

Figure 3.1 (a) Current distribution in Spain of <i>Pinus sylvestris</i> , <i>Pinus halepensis</i> and <i>Quercus ilex</i> according to the Spanish Forest Inventory (SFI) (b) Hypothesis diagram.....	63
Figure 3.2 Relationship between growth (NPP) and suitability (NBM weighted ensemble) for (a) <i>Quercus ilex</i> , (b) <i>Pinus sylvestris</i> and (c) <i>Pinus halepensis</i>	69
Figure 3.3 Positive and negative incongruence identified for <i>Pinus sylvestris</i> (a) and <i>Quercus ilex</i> (b) and their geographical location (c, d, respectively) along a suitability gradient (weighted ensemble of NBM)	70
Figure 3.4. Relationship of Leaf Area Index (LAI) and Mean Leaf Life (MLL) with suitability (NBM weighted ensemble) for <i>Quercus ilex</i> (a,b), <i>Pinus sylvestris</i> (c,d) and <i>Pinus halepensis</i> (e, f).....	73

CHAPTER 4

Figure 4.1 Current distribution of <i>Q.ilex</i> , <i>P.halepensis</i> and <i>P.sylvestris</i> in continental Spain	92
Figure 4.2 Variation of performance of each niche modeling technique.....	100
Figure 4.3 Changes in suitability in current forests	101
Figure 4. 4 Predicted future of current forests of <i>Quercus ilex</i> , <i>Pinus halepensis</i> , and <i>P. sylvestris</i> in continental Spain, as predicted by the multi model ensemble for the period 2050–2080	101

Figure 4.5 The spatially explicit change (percentage per pixel) in average per period net primary production (NPP) (GOTILWA1) and estimated Suitability (multi-niche-based model ensemble), between the periods 1950–1980 and 2050–2080, considering both GOTILWA+ simulations with (gray) and without (black) an atmospheric CO₂ increment. Lines represent linear regressions103

Figure 4.6 The spatially explicit change (percentage per pixel) in average per period net primary production (NPP) (GOTILWA+) and estimated Suitability (multi-niche-based model ensemble), between the periods 1950–1980 and 2050–2080, considering both GOTILWA+ simulations with (gray) and without (black) an atmospheric CO₂ increment104

CHAPTER 5

Figure 5.1 Framework of the niche concept.126

Figure 5.2 Spanish national forest inventory plot location of the co-occurrence of *Abies alba* and *Fagus sylvatica* (A), of *Abies alba* (B) and of *Fagus sylvatica* (C) in the Pyrenees and in the Iberian Peninsula127

Figure 5.3 Current and future suitability for *Abies alba* and *Fagus sylvatica* separately132

Figure 5.4 Agreement and disagreement between modeling Community and Overlapped Individual Models approach (CM and OIM) for present climate and future scenarios134

CHAPTER 6

Figure 6.1. Distribution of *Rhaponticum repens* (L.) in different realms of the world (different shaded areas)..152

Figure 6.2. Sequence of methods used.....156

Figure 6.3. Results of projections to the Iberian Peninsula, according to different provenance models.....157

Figure 6.4. Spatial projections of invasions of *Rhaponticum repens* in the Iberian Peninsula, according to different invasion provenance models.....159

CHAPTER 7

Figure 7.1 The bioclimatic velocity of climate change for different California endemic tree species within their range. (a) Bioclimatic velocity for the period present (1971-2000) to mid-century (2041-2070). (b) Temporal categorization bioclimatic velocity, whether higher in mid-century (2041-2070) or end of century (2071-2100)178

Figure 7.2 Climate change exposure metrics in species ranges in to period 1: present (1971-2000) to mid-century (2041-2070) and period 2: present (1971-2000) to end of century (2041-2070)179

Figure 7.3 Spatial configuration dynamics in species ranges in period 1: present (1971-2000) to mid-century (2041-2070); and period 2: present (1971-2000) to end of century180

Figure 7.4 Velocity of climate change exposure on current distribution in period 1: present (1971-2000) to mid-century (2041-2070); and period 2: present (1971-2000) to end of century181

Figure 7.5 Species' exposure to climate change for different dimensions in period 1: present (1971-2000) to mid-century (2041-2070); and period 2: present (1971-2000) to end of century182

INDEX OF TABLES

CHAPTER 3

Table 3.1 NBM used in this study	65
Table 3.2 Summary of incongruence values for each species considered and the associated environmental variables	71
Table 3.3 Correlation between modeled NPP and suitability for different niche-based model techniques	74

CHAPTER 4

Table 4.1 Niche-based models used in this study	95
Table 4.2 Assessment of the agreement between modeled and observed distributions for each niche-based model and species	99
Table 4.3 Percentage (%) of current forest stands which were predicted to become unsuitable for their current species by the period 2050–2080 according to the different statistical models	102

CHAPTER 5

Table 5.1 Summary of the GLMs	129
Table 5.2 Total suitable area of <i>Abies alba</i> and <i>Fagus sylvatica</i> obtained from present and two future (A1FI and A2) climatic scenarios	132
Table 5.3 Differences in suitable area and predicted changes for each modeling approach: OIM (Overlapped Individual Models) and CM (Community model).....	133
Table 5.4 Agreement between the community model and overlay individual model for the climate change scenarios considered.....	135
Table 5.5 Percentage of predicted suitable/non-suitable topo-climatic conditions for current NFI plots of mixed forest (<i>Abies alba</i> - <i>Fagus sylvatica</i>) for the two model approaches.....	136

CHAPTER 6

Table 6.1. Provenance Maxent models' accuracy, using the area under the receiving operating curve (ROC, (Fielding and Bell 1997))158

Table 6.2. Niche overlap of provenance models in the Iberian Peninsula measured by relativerank (RR), Schoener index (D) and modified Hellinger's distance (I)161

INDEX OF ACRONYMS

AIC Akaike information criterion	RN realized niche
AOGCM atmosphere-ocean general circulation models	ROC receiver operating characteristics
AUC area under de curve	RSA ranges spatial aggregation
BA basal area	RSF range suitability in forests
CM comunity model	SDM species distribution model
CRU climate research unit	SRC species range change
DCAIP digital climatic atlas of the Iberian Peninsula	SRE surface range envelope
DEM digital elevation model	TPR true positive rate
EET equate entropy threshold	TSS true skill statistic
ENM ecological niche models	VF virtual forest
FCE forest change envelope	
FCV forest change velocity	
FDA flexible discriminant analysis	
FME forest migration effort	
FN fundamental niche	
GAM generalized additive models	
GBIF global biodiversity information facility	
GCM global circulation models	
GFDL geophysical fluid dynamic laboratory	
GIS geographic information systems	
GLM generalized linear models	
GOTILWA growth is limited by water (model name)	
GPP gross primary production	
HadCM3 Hadley Centre Coupled Model, v.3	
IPCC intergovernmental panel for climate change	
LAI leaf area index	
MARS multiadaptive regression splines	
MAXENT maximum entropy	
MLL mean leaf life	
MTP minimun training presence	
NFI national forest inventory	
NPP net primary production	
OGC open geospatial consortium	
OIM overlap individual model	
PCA principal component analysis	
PCM parallel climate model	
RCV range change velocity	
REM range effort migration	
RMSE root mean square error	

1. INTRODUCTION

1.1 Preliminaries

Knowing, understanding and predicting species distributions under global change: A cross-disciplinary challenge.

Why do we want to know species distributions? What causes the observed pattern? I first stumbled on these questions years ago, during a field trip to the Pyrenees with Professor JM Roure in my last year of Environmental Science program. Since then, I have encountered many non-conclusive answers to them but I have to confess that back then, for an environmental scientist mainly trained in social sciences, there was not an obvious answer. I contented myself answering that understanding the ecology of the species was an aim itself and that enabled us to know the Earth system better, which I think it is still a valid statement nowadays.

Following my interest in environmental (*sensu lato*) controls on biota, Professor JM Roure pointed me to my PhD supervisor, Professor Ninyerola. He has been guiding me these years in my discovery of GIS science to undertake my first study on climatic classification (Serra-Diaz et al., 2011) and subsequent bioclimatic profiling of species through statistical modeling, which is being presented in this PhD dissertation.

After reading for a long while, I rapidly got to realize that understanding species distribution was a task far from trivial, and that many factors intertwine to produce what we see today: a certain species in a given geographical location. The role of both the environmental history of the Earth and the current environment is crucial to understand and predict species distributions: Past climatic changes (e.g. last glacial maximum), former and current uses of land by humans (e.g. treelines), environmental factors that affect species physiology (e.g. temperature tolerance), disturbance events (e.g. fire or droughts), species' use by humans (e.g. plantations), and so on. Largely

influenced by my colleagues in Geography (especially Dr. Pèlachs), I certainly broadened the scope and realized that species (and landscapes) are a clear footprint of the intersection between past and present.

Given this multiplicity of factors it did not surprise me so much the fact that many disciplines (and sub-disciplines) approached species distributions studies. Terms like chorology, biogeography, physical geography or ecology did not sound unfamiliar to me at all. However, I felt that recently a whole new terminology in (sub?) disciplines is being used to study species distributions: ecography, macroecology, spatial ecology, global biogeography. These new words rather emphasize details on the ecology of biogeography, or the biogeography of ecology. In my opinion, they represent a difficult conceptual mess, particularly for a newbie, even though I accept and sometimes embrace the value and subtleties of these terms.

Currently, the convergence of biogeography and ecology is a topic of theoretical discussion (Jenkins & Ricklefs, 2011; Dawson & Hortal, 2012). These authors identified a gradual merging of these scientific communities, particularly at the intermediate temporal and spatial scales. This example illustrates how disciplines are not static boxes and, like many processes in the real world, disciplines are dynamic: they split and merge and the borders between them tend to be rather fuzzy than sharp. Just the same story as it should be for university departments and research networks.

Each discipline and literatures used different conceptual models to understand and project the distributions of species, species richness and communities. Likewise, concepts like potential natural vegetation, niche, biotope or derivations of neutral theory of ecology have their own explanation (or not) of the patterns observed in the geography of species and biodiversity. I then understood that the concepts one chooses to approach species distributions may be contingent to one's background, but many concepts overlap.

In the field of botany, the term potential natural vegetation has been traditionally used to describe the state of mature vegetation in absence of human intervention. Such approach, traditionally based on phytosociological studies still brings some heated debates in the literature

(Carrión & Fernández, 2009; Chiarucci et al., 2010; Carrión, 2010; Loidi & Fernández - González, 2012 ; Somodi et al., 2012), mainly due to the ambiguous use of the term and its disputable operational value. However, this term is being applied in some instances in the modeling scene (Hickler et al., 2012).

Despite theoretical tribulations, I set my feet down to earth and I could not escape the fact that we are facing an era of global change, among which climate change has been one of the most studied phenomenon. In this context, I wondered again about the relevance of predicting future species distribution: Relocation of species in the future or their eventual extinction will likely affect human beings through the redistribution of ecosystem services and potential loss of ecological traditional knowledge, to mention a few. Thereby, I considered that studying future distributional shifts of species (and ultimately ecosystems and biomes) will enable us to adapt to the forthcoming environmental conditions, or mitigate its impacts. Overall, addressing species distribution represents a phenomenon of paramount interest for both biological research and management needs in the face of global change.

Many definitions for global change are out there, but I especially like the one coined by the U.S. Global Change Research Act of 1990: *“Changes in the global environment (including alterations in climate, land productivity, oceans or other water resources, atmospheric chemistry, and ecological systems) that may alter the capacity of the Earth to sustain life”*. Although it hides who to blame for this effects, the definition embraces the much of the high dimensionality of the issue. However, it lacks of accounting for changes in social environment, which is crucial also for predicting future scenarios. Global change is therefore as complex as species distributions, potentially hiding interacting factors yet to be identified.

In order to develop predictions of future distributional shifts I dove into the realm of modeling. Initially, the idea was to move beyond from the expert-based system of phytosociological studies, which I acknowledge its strengths, to a more statistical oriented approach capable of applying the same set of rules to the study area. In general, models provide a powerful and practical tool to study a system or a guide to decision-making processes and may be less judgmental. Following my initial training on GIS and statistics, we developed an atlas of potential distributional shifts in

Iberian trees (Ninyerola et al., 2010) using a correlative statistical model of environment with presence-absence of species. This first project constituted the cornerstone of the studies presented in this PhD thesis. Nearly all studies in this document deal with climate change predictions, which I acknowledge is only a subset of the wholesale of processes compounding global change. I believe this is a start, although I would have enjoyed including many other processes. Definitely, separating a PhD project with one's scientific career project is definitely a lesson I have learned.

In the beginning, I used the term "species distribution model" to refer to the correlative statistical approach of linking presence (optionally absence) of a species with environmental variables. However, during this PhD I have been changing the term throughout, at the risk of reader's annoyance. Many terms have been coined to this kind of models (environmental niche models, habitat suitability model, ecological niche models, etc.). There may be good reasons for choosing one or another. For instance, authors like J. Franklin prefer the use of the term species distribution models because the response variable is current species presence, which is influenced by many factors besides niche dimensions (Franklin, 2010a). Others, like P. Peterson and colleagues (Peterson et al., 2011) prefer to use niche models, because our final aim is to estimate the environmental conditions where a species lives (e.g. niche) and we do not take into account other variables. Therefore, whether the emphasis is given to the current distributions or the environmental factors may be a suitable criterion to use one term or another (Saupe et al., 2012).

However, in the introduction and final remarks of this PhD I will use the term ecological niche models (ENM), although in each chapter I may use a different term when current distribution or niche dimensions where more appropriate for the context. The reason why I chose ENM is based on the terminological classification developed by Sillero (2011). It enables the use of the term for both correlative and mechanistic models and directly links the statistical formulation to the theoretical background of the model. In this PhD thesis we will only use correlative ENM, so the acronym CENM will be used.

Finally, as scientists, we are engaged on identifying risks and finding solutions that forcefully imply the understanding of the Earth. As I am writing these words, new studies are warning that we are approximating to a (maybe) transition state in the Earth system (Barnosky et al., 2012), and

how society engagement is necessary to turn the tables of Earth's overexploitation (Ehrlich et al., 2012). These days, 20 years after the Conference of Rio (1992) for Biodiversity, countries have shown little success besides those beautiful but uncompromising words written in a new treatise. However I think that we sometimes need to rely on the naïve hypothesis that makes us (scientists) believe that setting better predictions of the geography of species will, ultimately, enable to both realize the magnitude of the issues we are facing and develop strategies for nature conservation in the Anthropocene era.

1.2 Setting the scene of species distribution models: a very brief summary (or not).

In the present work we will refer to correlative ecological niche models (CENM) as the phenomenological model using statistical techniques that enables characterizing and mapping species distributional areas. Such statistical models use georeferenced data on species distribution (presence and/or absence) and the environmental variables thought (or known) to influence (or determine) species requirements. Subsequently, a statistical technique is employed to fit the relationship between species occurrence and the environment it inhabits. After assessing model accuracy, the environmental-occurrence relationship is projected to the geographical space and distributional areas are mapped (see Figure 1.1).

The above definition hides the challenges that make the art of modeling species distribution more than a simple statistical juggling. We will revise below different options we can choose (or not) and outline several issues raised in the literature at each modeling step (Figure 1.1). It is not our intention here to write an extensive review of the current state of the art, for which we encourage reading (Guisan & Thuiller, 2005; Elith & Leathwick, 2009; Franklin, 2010a; Elith et al., 2010; Peterson et al., 2011). Instead, our goal is to summarize current topics in CENM in order to frame the context of the complexities contingent to modeling species distributions. The reader can also find interesting Appendix 1, in which an extended literature review is organized for different steps of the modeling .

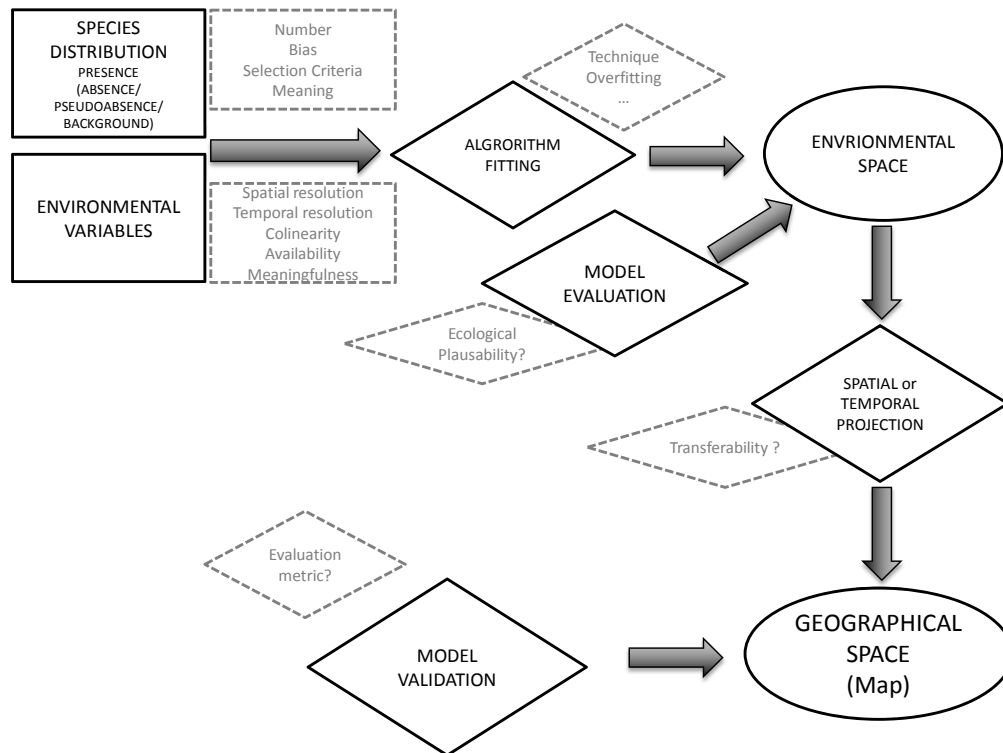


Figure 1.1 Modeling sequence of correlative ecological niche models.

What drives species distribution? Not so easy choosing predictors

Selecting what factors influence species distribution may not be obvious in many cases. Often, (bio)climatic variables are selected because analysis of species distributions are usually realized at coarse scales, however different environmental predictors are meaningful at different spatial and temporal scales. Austin (2002) and Pearson & Dawson (2003) showed different conceptual frameworks of the interaction of different predictors and scale with species distributions. In summary, while broad climate variables work at large scales, topography and soil work at finer spatial scales. Indeed, topographic predictors may increase model accuracy but they may be challenged by temporal transferability (discussed below), that is, model extrapolation to new climatic situations. For instance, it has been argued that using topography hinders CENM to be used for climate change scenarios, since some of the predictors remain constant. On the other hand, such predictions are a surrogate of climate variables working at lower spatial resolution and could

therefore conjecture results from broad climatic datasets and improve model accuracy (Lassueur et al., 2006). In a recent paper Stanton et al. (2011) shows that including such static variables is however recommendable for future predictions, if we are certain that it clearly affects distributions of target species. In general, selecting relevant variables in accordance with species ecology is not always straightforward as such information may be lacking for many species.

Parallel to this debate is the use of remote sensing variables in CENM (Kerr & Ostrovsky, 2003). The products from satellite imagery have been to a lesser extent applied for different purposes: identifying niche similarity to better predict rare plants and invasions (Roura-Pascual et al., 2006; Zimmermann et al., 2007; Papes et al., 2012) and even direct reflectance values have been incorporated to models (Morán-Ordóñez et al., 2012). The difficulties to model such predictions to the future, however, have hampered its widespread inclusion in CENM, whose scientific literature (not necessarily practice) has been very much focused on climate change effects.

There is a whole literature about the role of landscape and land use variables in these models although unfortunately they may have not reached the popularity or availability of bioclimatic parameters. For some taxa, especially birds and vertebrates, land-use variables and its influence in modeling outcomes have been widely discussed (Thuiller et al., 2004a; Austin & Van Niel, 2011; Triviño et al., 2011; Barbet-Massin et al., 2012). Additionally, the development of land use change models has boosted conservation studies that balance both climate and habitat loss in short-term global change studies (Bomhard et al., 2005; Araújo et al. 2008).

In any case, many of these variables may show a high degree of collinearity which may affect parameter estimation, especially when using regression techniques as fitting algorithm, as well as modeling transferability and interpretability (e.g. equifinality). Although many methods could be applied (Dormann et al., 2012a), common uses set a threshold to consider a non-collinear variable.

What is my data like? Choosing or adapting to data approaches and sampling strategies

Depending on the species data used for model building we can find 4 different approaches: (1) presence-only, (2) presence-pseudoabsence, (3) presence-background and (4) presence-absence. (1) Presence-only models profile the environmental conditions in which the species occur and are mostly used in the case of herbarium and museum data collections. (2) Presence-pseudoabsence is similar to the presence-only model but absence selection is not based on any record of absence but on a thought or random selection of probable absences. (3) Presence-background models use a set of locations in order to characterize species environment. The key difference is that background locations could coincide with occurrence locations whereas pseudo-absences do not. The rationale underneath is that pseudo-absences aim at simulating absence locations whereas background aims at environmentally characterizing the area under study. In practice, the species modeled normally constitute a small proportion of the geographical area under study and differences between background and pseudoabsence tend to be subtle (Peterson et al., 2011). Finally, (4) presence-absence models additionally use information of locations where the species is known not to be present and are used in the case of inventory data, where a designed sampling has been undertaken.

In selecting species geographical distribution, bias in presence data and pseudoabsence or background selection constitute one of the major issues. It has been proposed that, at least for presence-absence, background or pseudo-absence, a possible approach is to cancel bias of presences by reproducing bias in absences, but different techniques to perform such task have not been widely tested (but see (Phillips et al., 2009)). In addition, absence selection and the biological meaning of such absence play an influential role in model building (Chefaoui & Lobo, 2008; VanDerWal et al., 2009b; Lobo et al., 2010), essentially because we do not know the cause of such absence (e.g., humans? climate? species competition?). This may hinder true responses of species occurrence to the selected predictor variables. Furthermore, selecting the regions for background is not obvious and may rely on different assumptions that modelers need to deal with (Barve et al. 2011). However, some practical approaches have started to sprout out: from modeling potential

areas visited by the species (Barve et al. 2011) to the use of trend surface analysis (Acevedo et al. 2012).

Other sources of bias are detectability, locational errors and imperfect reference data, which have been less explored 'by default' in CENM studies. Indeed the first rule of a naturalist in the field is that some (sometimes most) of the species are undetected. The treatment of detectability issues has been addressed by some authors (Gu & Swihart, 2004; Royle et al., 2005; Hartel et al., 2009; Rota et al., 2011), who pointed out that detectability clearly influence parameter estimation and accuracy, especially in the case of logistic regressions. Georeferencing errors also decrease model performance (Graham et al., 2008; Osborne & Leitão, 2009) although CENM may retrieve acceptable accuracy under medium-error levels, specially machine-learning techniques (Graham et al., 2008). However, even having ecologically plausible and statistically meaningful models, Osborne et al. (2009) warns that variable selection and importance is widely affected by such positional errors and may therefore compromise inference.

Other data features that affect model accuracy are quantity and quality. I believe that the amount of data used have been widely studied (Wisz et al., 2008) compared to the quality of data (Dormann et al., 2008). A thorough analysis on imperfect reference data realized by Foody (2011) highlight the need for most of presence-absence to undertake a careful design, as the quality and quantity of data should meet the specific need of the study and may show other behavior besides changes in performance metrics. Last but not least, the relative abundant sources of data on species distributions has led to the use of different data sources, underlying different sampling strategies which clearly may affect our estimations in several ways (see Edwards et al. (2006) for examples of accuracy change).

Embracing a whole ecosystem of algorithms

These 4 approaches on geographic distribution of the target species, largely depending on available data and sampling method, bound the range of statistical techniques we can use, which could in turn be divided into 3 large (albeit fuzzy) groups: (1) regression techniques, (2) machine-learning techniques and (3) classification-similarity techniques (Franklin, 2010a). We could also

identify other largely used approaches, like expert-based models (Carter et al., 2006) and Bayesian approaches (Latimer et al., 2006), but they will not be used in this work.

Modern regressions have been widely applied and compared (Guisan et al., 2002; Segurado & Araujo, 2004), and they can use different type of functions to accommodate the shape of the response between independent variables and occurrence (dependent variable). Main algorithms employed are generalized linear models (GLM), generalized additive models (GAM) and multiadaptive regression splines (MARS). The difference mainly relies on the flexibility to accommodate response functions and whether they can or not apply interaction terms. Within regression models, spatial regressions deserve special consideration, because it has been largely recognized that undesired effects of spatial autocorrelation affect regression (Dormann, 2007), although other machine-learning methods too (Veloz, 2009). Methods like autologistic regression, generalized estimating equations, generalized mixed models or geographically weighted regressions have been, to a lesser extent, employed but they explicitly account for the spatial component (Dormann et al., 2007).

Machine learning algorithms imply to the detection of a set o rules that can classify input data accordingly; hence they are typically data-driven approaches. In the arena of machine learning, decision-tree methods have been widely applied and currently variations of them (e.g boosted regression trees, random forests) that are computationally more intensive and use a wide number of decision trees have proved to retrieve very good results (Elith et al., 2006). Other methods used in CENM are artificial neural networks, which classify data based on combinations of parameters and genetic algorithms, which develop a set of probabilistic rules that are sequentially selected. A less known algorithm currently being implemented are supported vector machines, which defines the hyperplane that divides presences from absences (Guo et al., 2005), although they could be used also for presence-only (Drake et al., 2006). Perhaps the most popular method among machine learning techniques is the so-called maxent, which stands for maximum entropy algorithm used in species distributions (Phillips et al., 2006). Such method seeks to define the distribution of maximum entropy within defined constraints, in our case relative to current species distribution. It has been emphasized that such computational intensive algorithms may be prone to overfit data, sometimes through overparametrization (Warren & Seifert, 2011). Such situation may lead to

ecologically unrealistic response functions and a less general model. Different regularization terms, especially in maxent, have been proposed to soften the response functions (Elith et al., 2011) and some new methodologies are being implemented to choose a suitable number of predictors (Warren & Seifert, 2011).

Classification-similarity techniques are algorithms normally based on presence only data, describing species locations in a multivariate statistical distance framework. Among these, enveloping techniques imply the use of mathematical algorithms based on defining the environmental space among presences. They define an environmental region in which the species is present and can range from rectangular fitting in the environmental space (using minimums and maximums of environmental layers; e.g. BIOCLIM) to more complicated “shapes” like convex hulls (e.g. HABITAT). Other methods, such as ecological niche factor analysis (Hirzel et al., 2002) use Mahalanobis distances from averaged conditions in which the species occurs to define occurrence probabilities.

Selecting data approach and algorithm produced several multi-model comparisons in the scientific literature, aiming to unveil which technique and approach performs best (Brotons et al., 2004; Muñoz & Felicísimo, 2004; Elith et al., 2006). In fact, it is important to stress that the combination of techniques and data provide different insights into characterizing different kinds of distributional areas and even different probabilistic measurements (probability of occurrence, probability of habitat use, etc.). For instance, Jiménez-Valverde et al. (2008) showed a simple but clarifying diagram of the conceptual implications of such choices (see Figure 1.2). On one extreme these authors indicate that presence-only enveloping models tend to characterize potential distributional areas better (potential niche), whereas presence-absence machine learning methods tend to approach to current or realized distributional areas (realized niche, see section 1.3 for further details). These different characterizations will be suited to different study goals and will largely depend on the modeler. For instance, studies aiming at discovering new species populations may want to draw potential distributional areas maps, whereas studies aiming at monitoring current populations may want to have a map that tightly render current distribution area.

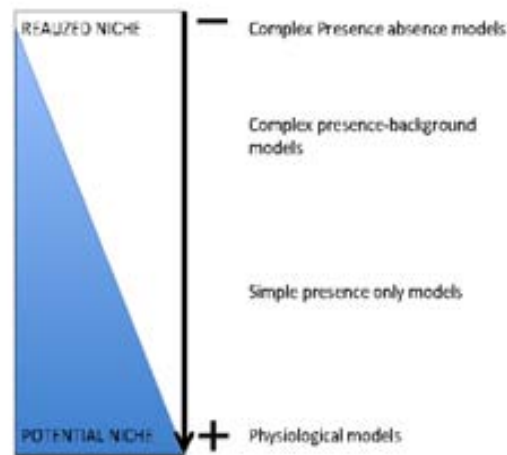


Figure 1.2 Relationship between algorithm and data approach with the type of niche being modeled. Modified from Jimenez-Valverde et al. (2008)

In order to reduce the uncertainty, which may be high especially (not exclusively) due to the algorithm employed, it has proposed the use of ensemble models (from the outcomes of several techniques) (Araújo & New, 2007; Rangel et al., 2009; Thuiller et al., 2009), which may increase overall accuracy results (Grenouillet et al., 2010). However, model ensembles can be constructed using different approaches (Araújo & New, 2007) and may not always increase the plausibility of the predictions (Elith et al., 2010; Mellert et al., 2011).

Spatial and temporal inference

After building the model, the fitted relationship between environmental variables and occurrences is applied to the landscape under analysis, the so-called model 'projection in space'. In such procedure every map unit is assigned to a 'probability of presence' derived from the built model. A key issue here is that the environmental range of model calibration or fitting should comprise the environmental range of projection; otherwise models are extrapolating beyond their model building capacities. Model extrapolation issues often arise not only when projecting to space

(i.e. different geographical areas) but also in time (i.e. future or past), when new environmental conditions will occur or have occurred (Williams & Jackson, 2007; Synes & Osborne, 2011; Zurell et al., 2012; Veloz et al., 2012). These extrapolation regions should be taken with caution when assessing past or future distributions. Some authors have applied distance operators in order to depict extrapolation regions in the geographical space (multivariate environmental similarity surface; Elith et al. 2010)

Projection capacity (in space or time) is also closely related to the selected algorithm. It has been outlined that model complexity, more variables and much complex response curves, may retrieve very accurate predictions in the geographical area of parameterization, but show little transferability to other regions (Heikkinen et al., 2011). Parameter settings are therefore crucial in order to avoid overfitting to data used to model building.

Not so obvious thresholds

In order to determine the area of distribution (whether potential or realized) it is necessary to convert probability of presence values into categories (suitable-presence vs. unsuitable-absence), therefore a threshold needs to be set. Researchers have nourished literature with different ways to do that (Liu et al., 2005; Jiménez-Valverde & Lobo, 2007) and many papers sprouted revealing major differences in modeling outcomes depending on the threshold chosen (Freeman & Moisen, 2008; Nenzén & Araújo, 2011). Generally, widely applied thresholds balance the correct classification of presences and absences (optimization thresholds: Maximum Kappa, Equate Sensitivity-Specificity) or maximize the correct classification of presences (i.e. sensitivity) or absences (i.e. specificity). However, optimization thresholds may not be the best approach when the costs of incorrectly predict presences and absences are different, which will be determined by the goal of the study. For instance, in the case of invasive species it is acknowledged that maximization of presences correctly predicted (true positive rate) constitute a better approach, because the costs of invasive species extraction surpass the costs of prevention (Jiménez-Valverde et al., 2011): underpredictions are more expensive than overpredictions in this case.

Is my model an accurate model?

Assessing performance and significance of CENM projections is not a trivial task and many measures may slightly indicate different characteristics of the power to predict species distributions (Jiménez-Valverde & Lobo 2007). In theory, independent datasets should be used for calibration/fitting and validation/evaluation, but in practice a common use is to partition data into training and testing according to a percentage that ranges from 70-80% to 30-20% respectively. However, in the recent years computational advances enabled the use of jackknifing and bootstrapping techniques. Such techniques use a high number of modeling iterations by partitioning data into n-subsets that will (with or without replacement of samples) test model accuracy repetitively. In principle, CENM and accuracy results should be much more robust using such techniques, because we ensure that all information on species occurrence has entered during model building.

Another issue is what metric should be used in order to assess model predictive ability and to my knowledge, here we are still far from consensus. Traditional approaches have broadly used different measures derived from a confusion matrix in which correctly and incorrectly presences or absences are quantified according to a certain chosen threshold. In the case of models using background information, modifications are undertaken so that instead of absences they calculate the area predicted to be suitable for the species, although other modifications of presence-absence metrics have been proposed (see Phillips et al. (2009); Peterson et al. (2008); Hirzel et al. (2006)). However, such approach largely relies on the threshold chosen which may hinder the evaluation of the model itself and potential comparison tests. Such situation has lead to the largely used threshold-independent measure of the area under the receiver operating characteristic curve (AUC; Fielding & Bell (1997)). AUC was initially developed in World War II and later applied to medical diagnostics (Thompson & Zucchini, 1989), it uses prediction errors across the whole range of predictions. Consequently, models and species can be compared because they are not threshold dependent. However, such measurement has received much criticism (Peterson et al., 2008; Lobo et al., 2008; Jiménez-Valverde, 2011). These authors point out many shortcomings of AUC such as:

low ability to account for the goodness of fit because it is just a discrimination measurement, poor discrimination in the central area of the curve and lacking information about spatial errors, among others.

Nevertheless, literature is still largely using AUC, probably because no better measurement has been proposed (but see modifications of AUC by Jiménez-Valverde & Lobo (2007)). To sum up, AUC should be accompanied by other threshold dependent techniques. In any case, accuracy measures should also be interpreted together with other parameters, because they largely interact with species traits and sample size (Hernandez et al., 2006; Syphard & Franklin, 2010; Syphard & Franklin, 2009). Despite intensive data-partitioning algorithms, we should always try to confront it to independent data and for this purpose, cross-scale data validation may be a good strategy to yield more reliable results (Marcer et al. 2012).

It is interesting to acknowledge here that model diagnostics in CENM have been mainly directed to discrimination measures. In other words, we assess how good models are on predicting suitable versus unsuitable conditions. Less explored measures of accuracy have analyzed the accuracy of the probability of presence, although some applications have rather used this range of values for subsequent analysis. In these cases a calibration curve (observed prevalence across the range of probability of presence) has been used to assess the goodness of the probability of presence (Acevedo et al., 2012). Ideally, in a perfect calibrated model, the probability of presence (from 0 to 1) and the observed species prevalence (ratio of predicted presences and total presences) should follow a 1:1 relationship (Franklin, 2010a).

Assuming a utopia?

Finally, it is important to recall that models represent a simplified understanding of a certain reality or process and therefore models rely on assumptions that may always be taken into account. Among others, most basic assumptions regarding CENM is that they assume that species are in equilibrium with the environmental variables selected (Araújo & Pearson, 2005) , we have a complete sampling of the species and knowledge of the whole set of key environmental variables controlling species distributions. In general, as mostly considered variables relate to climate and/or topography, it is often pointed out that biotic interactions and historical events (e.g. last glacial maximum) produce situations far from the equilibrium species-environment. Disequilibrium may vary among lineages and biogeographic regions (Munguía et al., 2012) and it is especially important when considering distributions of recent introductions, like in biological invasions (Václavík & Meentemeyer, 2012).

Another basic assumption is that populations and individuals of a species respond equally to environmental changes. Certainly, differences in environmental ranges have been found for biological subspecific entities (Thompson et al., 2011) and local adaptations may occur causing differences environmental requirements in the populations of the species under study (Benito Garzón et al., 2011). Overall, it may be discussed whether for some taxa; modeling at the species level may hinder global change predictions. Furthermore, the interaction of the species with the environment it inhabits is constant, although it is largely acknowledged processes like niche construction in some species (modification of the abiotic environment along individual or population development).

1.3 Theoretical framework: The history and (the continuous) evolution of the niche theory, and its relationship with species distribution.

Every model should be supported by an underlying existing theory and a set of assumptions. In our case, CENM are largely underpinned by niche theory, which has been one of the most difficult concepts to define in ecology (Chase & Leibold, 2003). In this section, we will revise early stages and current development of the theory and the niche concept, as well as outline potential connections to terminological issues found nowadays.

It is largely acknowledged that the niche concept is separately attributed to Joseph Grinnell (1917) and Charles Elton (1927). However, the first time the word 'niche' was published with an ecological interpretation was in 1910, by Johnson (1910). Grinnell first used the word niche in his PhD dissertation (1914) but he had previously used it in a publication by Grinnell & Swarth (1913). Therefore, it has been suggested that the use of 'niche' was already discussed among Stanford PhD students and used in an informal way (Wake et al., 2009). However, the first coining of the term is attributed to Grinnell, in a work where the word niche was explicitly stated: 'The niche relationship of the California thrasher', in 1917. The concept of 'niche' under Grinnell terms refers to the set of environmental conditions that determine the control on species distributions. Such definition was especially related to physiological tolerances that bound species ranges, especially temperature (Grinnell 1917b), although he was aware of the multifactorial nature of species distributions within these ranges.

Elton built a slightly different perspective of the same concept significantly later, in 1927, apparently independently from Grinnell (after Hutchinson and Elton correspondence). In his work 'Animal ecology' (Elton 1927), the use of niche is linked to a functional approach, as a role of a species in a community or in a food chain. Therefore, Elton's definition is based on biotic interactions whereas Grinnell's definition is more environmental/habitat oriented. Nevertheless,

some authors state that the main difference between such definitions relies, ultimately, from whether one or more species can occupy the same niche (Griesemeier (1992); although see other conceptual reviews in Udvardy (1959) and Hardin (1960)).

It is essential noting the working scales at which the two visions of the concept operate: Elton's vision of biotic interactions or impacts tends to be fine grained whereas Grinnell's vision based on resources typically works at larger scales (but see (Araújo & Luoto 2007) for importance of biotic interactions depicting large scaled patterns). Such different visions have encouraged the ENM literature to distinguish between Grinnellean and Eltonian niches (Hirzel & Lay 2008).

Certainly, the operational framework of the theory is attributed to George Evelyn Hutchinson (Hutchinson, 1957). In his famous 'Concluding remarks' in the Cold Spring Harbor Symposia on Quantitative Biology in 1957, he defined the niche as the N-dimensional 'hypervolume' of resources used by an organism. In addition, and influenced also by Elton's ideas, he further distinguished between the set of abiotic conditions in which a species can live ('fundamental niche') and the reduced set of conditions occupied by the species due to competition with other organisms mainly, but not exclusively ('realized niche'). These concepts are still widely present in the species distribution modeling literature, because species presence is actually a consequence of different biotic pressures; hence it has been argued that CENM model realized niches.

It is, however, much later (1978) that Hutchinson developed the ideas that have fed current 'niche literature' (Hutchinson 1978). From his conclusions, we can deduce that the key evolution of the concept is that its definition moved from a property of the physical environment to a property of the species (Colwell 1992), hence the core difference between the niche and the biotope (the physical space), which has been termed the 'Hutchinson's duality' (Colwell & Rangel, 2009). As we will see, such duality has encouraged new formulations of the niche dimensions.

Recently, the niche theory underpinning CENM has been interpreted in a more integrative approach, especially with population ecology. Pulliam (2000) integrated niche concept with metapopulation theory and source-sink theory in order to understand the link between species distributions and the different niches of a species (realized vs. potential). Accordingly, Jackson &

Overpeck (2000) pointed out a further link to geography, suggesting that most of Hutchinson's multidimensional niche was not actually appearing in our realized environment. They proposed the use of the term potential niche as the joint distribution of the realized environment and the realized niche. We argue that the raise of such intermingled terms has led to some confusion in literature and further hypothesize that different names of CENM may be a consequence of many interchangeable conceptualizations (e.g. species distribution models, environmental niche models, ecological niche models; Sillero (2011)).

Another important contribution to the ongoing understanding and evolution of the niche concept has been performed by Holt (Holt, 2009). He pointed out demographical features that potentially affect the realized environment of the species, such as density-dependent modulation and the role of dispersal and spatiotemporal dynamics of niche evolution. Altogether, such understanding challenges the traditional Hutchinsonian niche.

I believe that Jorge Soberón (2007) has successfully arranged a conceptual model that merges both niche theory with population dynamics in geographical space (see also contribution (Soberón & Nakamura, 2009). Soberon (2007) drew the BAM diagram (Figure 1.3), explaining that three dimensions that affect species distributions in a geographical space (G): abiotic factors (A), biotic factors (B) and available space (M). Species may be found in different combinations of subsets of these three dimensions always within the M dimension (space reachable by the species). However, only when the three dimensions are suitable we can identify source populations (positive growth rate) whereas others represent sink populations. Such conceptualization, albeit interesting, is practically ignored in many cases as most ecological knowledge on the species and the system is unfortunately largely lacking, especially concerning biotic dimension (B). Such situation has led to the formulation of the 'Eltonian noise hypothesis' (Peterson, 2005), that considers that most suitable abiotic conditions in a given geographical space are also suitable in the biotic dimensions, therefore biotic interactions can be interpreted as noise in data or unexplained variance. This hypothesis does not hold when large scale spatially distributed signals of biotic interaction appear, which may sometimes be the case (Hampe, 2004). However, it may be valid at broad scales using climatic predictors (Peterson et al., 2011).

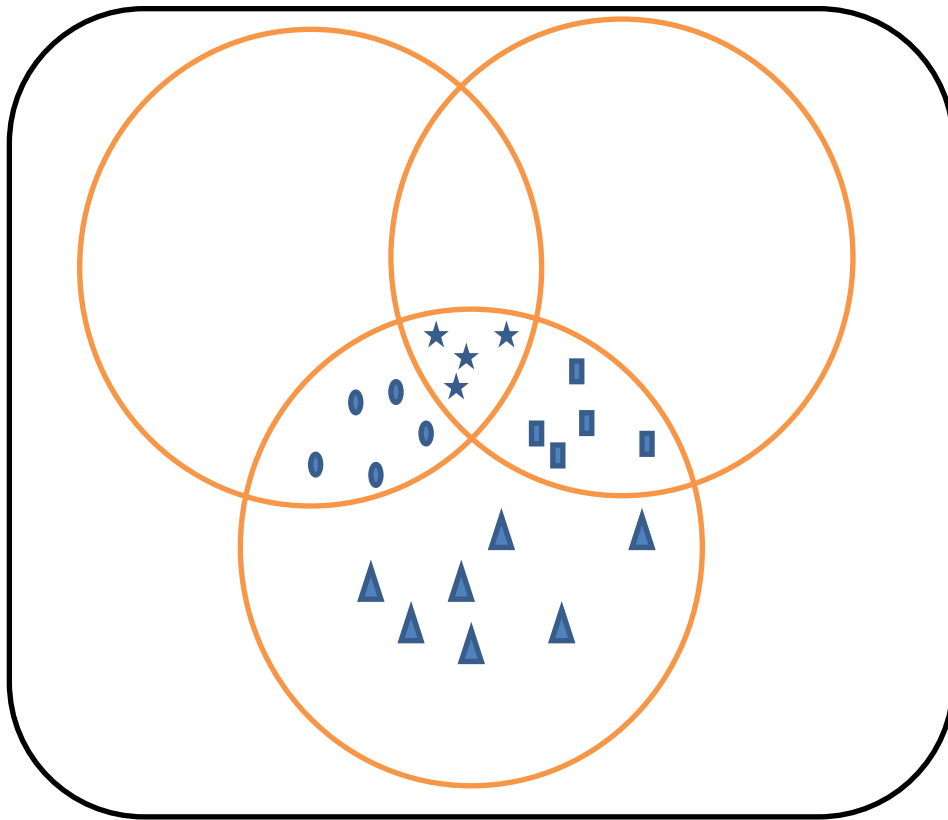


Figure 1.3 Conceptual framework of species distribution of the BAM diagram. G represents the geographical space under study, M represents the accessible geographical space by a species, A the abiotic conditions favorable for growth, reproduction and maintenance, B is the biotic dimension favorable to species persistence (e.g. it may either co-occur or exclude competitors). Star signs represent source populations, where the three dimensions intersect. Ellipsoids, rectangles and triangles represent sink populations where populations growth is negative due to biotic interaction (e.g. ellipsoids, outside B) or to abiotic conditions, (e.g. rectangles, out of A) or the combinations of the two of them (e.g. triangles, outside B and A)

1.4 Modeling framework

A model is a mathematical formulation (or a set of them) that aims at describing or simulating a certain real process or phenomenon. To accomplish that, many approaches have been used and many different categorizations of models have been proposed (e.g. Wissel, 1992).

Correlative ecological niche models (CENM) could be classified as a phenomenological or empirical model that aims at describing or predicting current species distributions. The phenomenon under analysis is species presence (sometimes abundance) and a set of predictors (abiotic or not) are used to describe what we empirically observe: current species presence.

A model typically built under such observations is inherently a static model, meaning that the state of the predictors does not change over time. Therefore, such models actually constitute a snapshot of the current state and no other transition states are considered under such modeling framework. Normally, variables in such models tend to be defined as averaged values over certain period of time. As a consequence, time is forgotten and a steady-state of the ecosystem is assumed. Several advantages may be typically characteristic of static models: less computationally intensive, fewer and easier parameterization, etc. Jørgensen & Bendoricchio (2001) highlight that static models constitute good pictures of average situations and large elements of the system may be included, however they also warn that because time is not described, transitions are not accounted for and extrapolation to other systems should not be valid.

If we consider that "essentially all models are wrong" (Box, 1987), there is no kind of model that is able to satisfy the three desired model intrinsic properties at the same time: Precision, Generality and Reality (Levins, 1966; Sharpe, 1990; Figure 1.4). Indeed, it is difficult to put together highly precise models that may be globally applicable and trustworthy describe reality. However, such conceptual scheme of trade-offs has been criticized by several authors (e.g. Korzukhin et al., 1996) who argue that such features may not necessarily be mutually exclusive. Nevertheless, we concur with (Guisan & Zimmermann, 2000) that it is a useful framework to approach a certain modeling

project, whose outcomes will be needed to be more general, precise or “real” depending on the model’s goal.

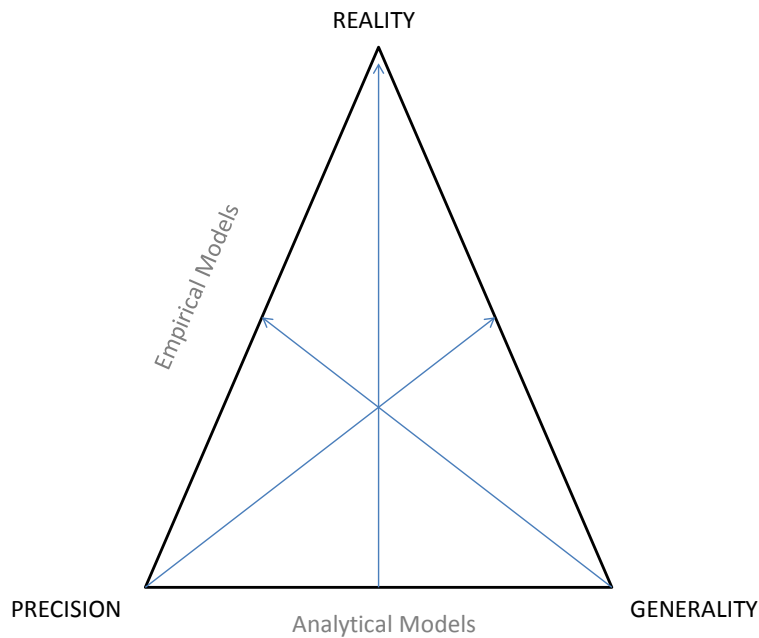


Figure 1.4 Model classification between trade-off scheme of 3 properties: Precision, generality and reality. After Levins (1966) and inspired by Guisan and Zimmermann (2000). Arrows indicate the furthest property for a given model

CENM may fall in the classification of precise and real models (Figure 1.4), but they may lack general application because we actually do not model the mechanism underlying the causes (climate, competition, etc.) and effects (species presence). As we will see in following sections, such distinctions have favored a great debate among species distribution modelers between correlative approaches, like CENM, with mechanistic approaches to modeling species distributions (Kearney, 2006; Kearney & Porter, 2009; Dormann et al., 2012b).

1.5 Applied framework of CENM in the face of global change

Despite aired critiques to CENM in ecological management (Sinclair et al., 2010) the fact is that applications of CENM continue to increase and many managers are still referring to their outcomes to orient many policies (Iverson et al., 2011). We argue that CENM still yield useful results in management and applied sciences when implemented with full understanding of assumptions and when critically interpreting results. We would like to point out that combination of current expert knowledge of species and CENM may picture a realistic scenario of the exposure of species (see Matthews et al. (2011) for an application on Eastern US forests).

I would like to outline here several CENM applications that have fruitfully yielded positive results. The great number of applications reveals that species distribution models and species distribution in general have a great importance in the field of applied and theoretical sciences, although here we will only succinctly point out those related to global change applications. A complete review may be found in Franklin (2010b) and Peterson et al. (2011).

Biodiversity monitoring and discovering

CENM can inform about potential suitable areas for a species, which have clear implications for discovering new populations, especially for rare and cryptic species. This has the potential to detect not sampled regions with similar environmental conditions as in the already sampled populations. There is a general agreement that we are undergoing the 6th mass extinction (Barnosky et al., 2011), therefore efficiently sampling geographical areas to find new populations of endangered or considered extinct populations is of paramount interest.

Using CENM, Guisan et al. (2006) tested a sequential strategy of modeling –sampling, resulting in an increased discovery of populations of an alpine plant *Eryngium alpinum*, and improved sampling efficiency by two times. Bourg et al. (2005) discovered 8 populations of a rare plant (*Xerophyllum asphodeloides*) using niche modeling based on classification and regression trees. De Siqueira et al. (2009) used landscape similarity analysis and found 6 additional populations of

Brazilian cerrado (*Byrsonimia subterranean*). Interestingly, Williams et al. (2009) found that machine learning algorithms performed better than regression based algorithms to find new populations of a rare edaphic specialist plant. The use of ensemble models and different resolutions can increase the chance of finding rare species: Le Lay et al. (2010) increased up to 50% the number of records in rare plants of the Swiss alps.

Surprisingly, these applications have even resulted in the discovery of new species: Raxworthy et al. (2003) developed different models for species of the same genera, and results of post-sampling ended up with the discovery of three new species of chameleon. Therefore, if niche conservatism holds for a group of selected species (e.g. related species tend to conserve their environmental niche Wiens et al. (2010); Peterson (2011)), we could expect the discovery of phylogenetically related species.

Conservation planning and management

CENM have been extensively applied in biological conservation science for many purposes. I argue that it has been of special importance in the spatial dimension of conservation, where the question 'where' to conserve has had clear implications for prioritizing hotspots of biodiversity (Mittermeier et al., 1998; Myers et al., 2000). Additionally, CENM could constitute a key step in planning species reintroductions (Hirzel et al., 2004; Martinez-Meyer et al., 2006).

Most applications to global change projections use CENM to account for changes in species ranges under future global warming scenarios (e.g. Thuiller et al. (2005); Benito Garzón et al. (2008)), and to a lesser extent, they even consider projections in land use changes (Barbet-Massin et al. 2012). Derived changes in species ranges have been used to map potential consequences for current reserve planning and future mitigation strategies (Araújo et al., 2011). However, the rationale used is that changes in species range are a good surrogate for extinction risk (Thomas et al., 2004), which is not necessarily the case (Thuiller et al., 2004b). Indeed extinction is derived from many processes among which area, but also stochastic processes play an important role. However, some comparisons yielded good results confronting CENM and other measures related to extinction (Araújo et al., 2002; VanDerWal et al., 2009a). Current approaches consist of linking

population ecology with results derived from CENM (Keith et al., 2008; Anderson et al., 2009; Brook et al., 2009; Regan et al., 2011).

Species invasions

One of the most global threats to ecosystems constitutes spread of invasive species globally (Millenium Ecosystem Assessment, 2005). CENM in this case have been widely applied to predict potential invasion risk in future conditions and in different areas. The idea behind is that we can capture the environmental suitable space of the invasive species and project it to our assessment area. Such approach has produced good results (Roura-Pascual et al., 2004; Peterson, 2005; Drake & Lodge, 2006), but equilibrium assumptions of CENM clearly challenge disequilibrium situations with climate in invasions, at least in the initial phases, as well as potentially different biotic interactions in native and different invaded ranges.

Currently, invasive CENM approaches tend to couple invasions with demographic processes as well as incorporating introduction efforts derived from world trade routes (Herborg et al., 2007). This emphasizes the broadening of the scope CENM to include other parameters besides climate (Gallien et al., 2010; Roura-Pascual et al., 2011)

Epidemiology and disease transmission studies

Although it has not been profoundly explored, several attempts have been undertaken to apply CENM to identify potential changes in distribution of disease vectors and reservoirs (Peterson & Shaw, 2003; Ogden et al., 2006; Nakazawa et al., 2007; Reed et al., 2008). Normally, pathogen causing diseases implies several vectors and reservoirs interacting at different spatial and temporal scales (e.g. ephemeral distributions). Therefore, it is clear that mapping disease risk under global change predictions requires the additional challenge of biotic interactions and temporal resolutions.

To our knowledge, some authors have used a similar approach to species, by modeling disease occurrences to environmental predictors and mapping the 'realized niche' of the disease (Yeshiwondim et al., 2009). Although useful, such approach may be biased to existing public health data or available ecotoxicological studies and results might be of little transferability.

Peterson (2009) analyzed future vectors of malaria transmissions in Africa over scenarios of climate change and reported significant poleward distributions. González et al. (2010) recently warned that leishmaniasis disease could potentially increase its area of influence in the USA due to northward expansions, and suggest potential control of reservoirs. Although such approach has been gaining popularity, still little evidence is found to earlier projections of infectious diseases spread after a century of global warming (Lafferty 2009), which suggests that the spread of vectors and reservoirs of transmission may be constrained by other factors than climate.

1.6 Aims and scope

The aim of the present work is to assess the usefulness of correlative ecological niche models (CENM) and propose better strategies to predict future species distributions. Our context and derived hypothesis will be framed in the ongoing global change conditions, especially climate change which poses many challenges to the application of these models.

The present work does not intend to present an extended collection of studies dealing with each identified issue in CENM, but rather emphasize some key processes that challenge the biological meaning of these models.

Taking previous identified gaps as a starting point, we developed specific studies addressing:

Species potential versus realized niche discrepancies.

This is a key analysis since we should determine to what extent climate or other factors are driving species distributions. Such analysis enhance the understanding of our target species.

The role of biogeochemistry under future global warming scenarios.

Traditionally estimations of climate change have been used to analyze changes in distributions due to eventual increase in temperature and eventual decreases in water availability. However, for plants, the role of CO₂ has been neglected and may affect performance of plants under global warming scenarios.

Differences between species individual vs. multi-species models.

In order to protect ecosystems many models have been calibrated at the species level. However, others advocate for using a broader biological level of organization (communities or ecosystems) to model species. It is important to assess how predictions may change between these two approaches.

Potential changes of species' niche of invasive species across geographical regions.

At present, species is used as a whole. However, different studies point out niche changes between populations, especially in invasions. Therefore, it is important to assess how these predictions may change and identify potential solutions.

Incorporating the temporal dimension in CENM predictions of exposure.

The static nature of CENM models have been widely criticized. It is important to advance in the potential dynamic outcomes of CENM in order to take the temporal dimension into account.

1.7 References

- Acevedo P, Jiménez-Valverde A, Lobo JM, Real R (2012) Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*, in press.
- Anderson BJ, Akçakaya HR, Araújo MB, Fordham DA, Martinez-Meyer E, Thuiller W, Brook BW (2009) Dynamics of range margins for metapopulations under climate change. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1415-1420.
- Araújo MB, Alagador D, Cabeza M, Nogués - Bravo D, Thuiller W (2011) Climate change threatens European conservation areas. *Ecology Letters*, **14**, 484-492.
- Araújo MB, Williams PH, Fuller RJ (2002) Dynamics of extinction and the selection of nature reserves. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **269**, 1971-1980.
- Araújo MB, Nogués-Bravo D, Reginster I, Rounsevell M, Whittaker RJ (2008) Exposure of European biodiversity to changes in human-induced pressures. *Environmental Science & Policy*, **11**, 38-45.
- Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, **16**, 743-753.
- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42-47.
- Araújo MB, Pearson RG (2005) Equilibrium of species' distributions with climate. *Ecography*, **28**, 693-695.
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101-118.
- Austin MP, Van Niel KP (2011) Impact of landscape predictors on climate change modelling of species distributions: a case study with *Eucalyptus fastigata* in southern New South Wales, Australia. *Journal of Biogeography*, **38**, 9-19.
- Barbet-Massin M, Thuiller W, Jiguet F (2012) The fate of European breeding birds under climate, land-use and dispersal scenarios. *Global Change Biology*, **18**, 881-890.
- Barnosky AD, Hadly EA, Bascompte J et al. (2012) Approaching a state shift in Earth's biosphere. *Nature*, **486**, 52-58.

- Barnosky AD, Matzke N, Tomiya S et al. (2011) Has the Earth/'s sixth mass extinction already arrived? *Nature*, **471**, 51-57.
- Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J, Villalobos F (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **11**, 1810-1819.
- Benito Garzón M, Sánchez de Dios R, Sainz Ollero H (2008) Effects of climate change on the distribution of Iberian tree species. *Applied Vegetation Science*, **11**, 169-178.
- Benito Garzón M, Alía R, Robson TM, Zavala MA (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766-778.
- Bomhard B, Richardson DM, Donaldson JS et al. (2005) Potential impacts of future land use and climate change on the Red List status of the Proteaceae in the Cape Floristic Region, South Africa. *Global Change Biology*, **11**, 1452-1468.
- Bourg NA, McShea WJ, Gill DE (2005) Putting a CART before the search: successful habitat prediction for a rare forest herb. *Ecology*, **86**, 2793-2804.
- Box, George E. P.; Norman R. Draper (1987). *Empirical Model-Building and Response Surfaces*, Wiley, p. 424.
- Brook BW, Akçakaya HR, Keith DA, Mace GM, Pearson RG, Araújo MB (2009) Integrating bioclimate with population models to improve forecasts of species extinctions under climate change. *Biology Letters*, **5**, 723-725.
- Brotons L, Thuiller W, Araújo MB, Hirzel AH (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437-448.
- Carrión JS (2010) The concepts of potential natural vegetation (PNV) and other abstractions (trying to pick up fish with wet hands). *Journal of Biogeography*, **37**, 2213-2215.
- Carrión J,S., Fernández S (2009) The survival of the 'natural potential vegetation' concept (or the power of tradition). *Journal of Biogeography (J.Biogeogr.)*, **36**, 2202-2203.
- Carter GM, Stolen ED, Breininger DR (2006) A rapid approach to modeling species–habitat relationships. *Biological Conservation*, **127**, 237-244.

- Chase JM, Leibold MA (2003) *Ecological niches :linking classical and contemporary approaches*. University of Chicago Press, Chicago etc., 212 pp.
- Chefaoui RM, Lobo JM (2008) Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling*, **210**, 478-486.
- Chiarucci A, Araújo MB, Decocq G, Beierkuhnlein C, Fernández-Palacios JM (2010) The concept of potential natural vegetation: an epitaph? *Journal of Vegetation Science*, **21**, 1172-1178.
- Colwell RK (1992) In *Keywords in Evolutionary Biology*. Eds Keller EF, Lloyd EA Harvard University Press, Cambridge, MA. Pp.241–248.
- Colwell RK, Rangel TF (2009) Hutchinson’s duality: The once and future niche. *Proceedings of the National Academy of Science*, **106**, 19651–19658.
- Conference Rio de Janeiro for Biodiversity (1992) Report of the United Nations conference on environment and Development. <http://www.un.org/documents/ga/conf151/aconf15126-1annex1.htm>
- Dawson MN, Hortal J (2012) A cure for seeing double? Convergence and unification in biogeography and ecology. *Frontiers of Biogeography*, **4**, 3-6.
- De Siqueira MF, Durigan G, de Marco Júnior P, Peterson AT (2009) Something from nothing: using landscape similarity and ecological niche modeling to find rare plant species. *Journal for Nature Conservation*, **17**, 25-32.
- Dormann CF (2007) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Ecology and Biogeography*, **16**, 129-138.
- Dormann CF, Purschke O, Márquez JRG, Lautenbach S, Schröder B (2008) Components of uncertainty in species distribution analysis: a case study of the great grey shrike. *Ecology*, **89**, 3371-3386.
- Dormann CF, Elith J, Bacher S et al. (2012a) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, no-no.
- Dormann CF, Schymanski SJ, Cabral J et al. (2012b) Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, no-no
- Drake JM, Lodge DM (2006) Forecasting potential distributions of nonindigenous species with a genetic algorithm. *Fisheries*, **31**, 9-16.

- Drake JM, Randin C, Guisan A (2006) Modelling ecological niches with support vector machines. *Journal of Applied Ecology*, **43**, 424-432.
- Edwards J, Thomas C., Cutler DR, Zimmermann NE, Geiser L, Moisen GG (2006) Effects of sample survey design on the accuracy of classification tree models in species distribution models. *Ecological Modelling*, **199**, 132-141.
- Elith J, Graham CH, Anderson RP et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151.
- Elith J, Kearney M, Phillips S (2010) The art of modelling range - shifting species. *Methods in ecology and evolution*, **1**, 330-342.
- Elith J, Leathwick JR (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 SJ, Hastie T, Dudík M, Chee YE, Yates CJ (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43-57.
- Ehlich PR, Kareiva PM, Daily GC (2012) Securing natural capital and expanding equity to rescale civilization . *Nature*, **486**, 68-73.
- Elton CS (1927) *Animal Ecology*. Sidgwick and Jackson, London.
- Dormann C, M McPherson J, B Araújo M et al. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609-628.
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38.
- Foody GM (2011) Impacts of imperfect reference data on the apparent accuracy of species presence?absence models and their predictions. *Global Ecology and Biogeography*, **20**, 498-508.
- Franklin J (2010a) *Mapping species distributions :spatial inference and prediction*. Cambridge University Press, Cambridge, 320 pp.
- Franklin J (2010b) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, **16**, 321-330.

- Freeman EA, Moisen GG (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, **217**, 48-58.
- Gallien L, Münkemüller T, Albert CH, Boulangeat I, Thuiller W (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, **16**, 331-342.
- González C, Wang O, Strutz SE, González-Salazar C, Sánchez-Cordero V, Sarkar S (2010) Climate change and risk of leishmaniasis in North America: predictions from ecological niche models of vector and reservoir species. *PLoS neglected tropical diseases*, **4**, e585.
- Graham CH, Elith J, Hijmans RJ, Guisan A, Townsend Peterson A, Loisele BA, The Nceas Predicting Species Distributions Working Group (2008) The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, **45**, 239-247.
- Grenouillet G, Buisson L, Casajus N, Lek S (2010) Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography*, **34**, 9-17.
- Griesemer JR (1992) In *Keywords in evolutionary biology*. Eds Keller EF, Lloyd EA (Harvard UnivPress, Cambridge, MA), 231–240
- Grinnell J (1914) An account of the mammals and birds of the Lower Colorado Valley with special reference to the distributional problems presented. University California Publ Zoology, **12**, 51–294.
- Grinnell J, Swarth H (1913) An account of the birds and mammals of the San Jacinto area of Southern California. University California Publ Zoology, **10**, 197– 406.
- Grinnell J (1917) The niche-relationships of the California thrasher. *The Auk*, **34**, 427– 433.
- Grinnell J (1917b) Field tests of theories concerning distributional control. *American Naturalist*, **51**, 115–128.
- Gu W, Swihart RK (2004) Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. *Biological Conservation*, **116**, 195-203.
- Guisan A, Broennimann O, Engler R, Vust M, Yoccoz NG, Lehmann A, Zimmermann NE (2006) Using niche - based models to improve the sampling of rare species. *Conservation Biology*, **20**, 501-511.

- Guisan A, Edwards TC, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, **157**, 89-100.
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993-1009.
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147-186.
- Guo Q, Kelly M, Graham CH (2005) Support vector machines for predicting distribution of Sudden Oak Death in California. *Ecological Modelling*, **182**, 75-90.
- Hampe A (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography*, **13**, 469-471.
- Hardin G (1960) The competitive exclusion principle. *Science*, **131**, 1292–1297.
- Hartel T, Ollerer K, Farczády L, Monga C, Bancila R (2009) Using species detectability to infer distribution, habitat use and absence of a cryptic species: the smooth snake (*Coronella austriaca*) in Saxon Transylvania. *Acta Scientiarum Transylvanica*, **17**, 61-76.
- Heikkinen RK, Marmion M, Luoto M (2011) Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography*, **35**, 276-288.
- Herborg LM, Jerde CL, Lodge DM, Ruiz GM, MacIsaac HJ (2007) Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecological Applications*, **17**, 663-674.
- Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773-785.
- Hickler T, Vohland K, Feehan J et al. (2012) Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. *Global Ecology and Biogeography*, **21**, 50-63.
- Hirzel AH, Posse B, OGGIER PA, Crettenand Y, Glenz C, Arlettaz R (2004) Ecological requirements of reintroduced species and the implications for release policy: the case of the bearded vulture. *Journal of Applied Ecology*, **41**, 1103-1116.
- Hirzel A, 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 AH, Lay GL (2008) Habitat suitability modelling and niche theory. *Journal of Applied Ecology*, **45**, 1372-1381.
- Hirzel AH, 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.
- Holt RD (2009) Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences*, **106**, 19659-19665.
- Hutchinson GE (1957) Concluding remarks. *Cold Spring Harbor Symposia*, **22**, 415-427.
- Hutchinson GE (1978) *An Introduction to Population Ecology*. Yale University Press, New Haven, CT.
- Iverson LR, Prasad AM, Matthews SN, Peters MP (2011) Lessons learned while integrating habitat, dispersal, disturbance, and life-history traits into species habitat models under climate change. *Ecosystems*, **14**, 1005-1020.
- Jackson ST, Overpeck JT (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, **26**, 194-220.
- Jenkins DG, Ricklefs RE (2011) Biogeography and ecology: two views of one world. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **366**, 2331-2335.
- Jiménez-Valverde A, Lobo JM (2007) Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecologica*, **31**, 361-369.
- Jiménez-Valverde A, Peterson A, Soberón J, Overton J, Aragón P, Lobo J (2011) Use of niche models in invasive species risk assessments. *Biological Invasions*, **13**, 2785-2797.
- Jiménez-Valverde A (2011) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, **21**, 408-597.
- Jiménez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, **14**, 885-890.
- Johnson RH (1910) *Determinate Evolution in the Color Pattern of the Lady-Beetles*. Publication, **122**, (Carnegie Institute of Washington, Washington, DC).
- Jørgensen SE, Bendoricchio G (2001) *Fundamentals of ecological modelling*. Elsevier, Amsterdam; New York, 530 pp.

- Kearney M (2006) Habitat, environment and niche: what are we modelling? *Oikos*, **115**, 186-191.
- Kearney M, Porter W (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species? ranges. *Ecology Letters*, **12**, 334-350.
- Keith DA, Akçakaya HR, Thuiller W et al. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560-563.
- Kerr JT, Ostrovsky M (2003) From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution*, **18**, 299-305.
- Korzukhin MD, Ter-Mikaelian MT, Wagner RG (1996) Process versus empirical models: which approach for forest ecosystem management? *Canadian Journal of Forest Research*, **26**, 879–887.
- Lafferty KD (2009) The ecology of climate change and infectious diseases. *Ecology*, **90**, 888-900.
- Lassueur T, Joost S, Randin CF (2006) Very high resolution digital elevation models: Do they improve models of plant species distribution? *Ecological Modelling*, **198**, 139-153.
- Latimer AM, Wu S, Gelfand AE, Silander Jr JA (2006) Building statistical models to analyze species distributions. *Ecological Applications*, **16**, 33-50.
- Le Lay G, Engler R, Franc E, Guisan A (2010) Prospective sampling based on model ensembles improves the detection of rare species. *Ecography*, **33**, 1015-1027.
- Levins, R., 1966. The strategy of model building in population ecology. *American Scientist*, **54**, 421–431.
- Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385-393.
- Lobo JM, Jiménez-Valverde A, Hortal J (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103-114.
- Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145-151.

- Loidi J, Fernández - González F (2012) Potential natural vegetation: reburying or reborning? *Journal of Vegetation Science*, **23**, 596-604.
- Marcer A, Pino J, Pons X, Brotons L (2012) Modelling invasive alien species distributions from digital biodiversity atlases. Model upscaling as a means of reconciling data at different scales. *Diversity and Distributions*, in press.
- Martinez-Meyer E, Peterson AT, Servín JI, Kiff LF (2006) Ecological niche modelling and prioritizing areas for species reintroductions. *Oryx*, **40**, 411-418.
- Matthews SN, Iverson LR, Prasad AM, Peters MP, Rodewald PG (2011) Modifying climate change habitat models using tree species-specific assessments of model uncertainty and life history-factors. *Forest Ecology and Management*, **262**, 1460-1472 .
- Mellert K, Fensterer V, Küchenhoff H, Reger B, Kölling C, Klemmt H, Ewald J (2011) Hypothesis - driven species distribution models for tree species in the Bavarian Alps. *Journal of Vegetation Science*, **22**, 635-646.
- Millenium Ecosystem Assessment (2005) Millennium ecosystem assessment. Ecosystems and Human Well-being: Synthesis Washington DC: Island Press
- Mittermeier RA, Myers N, Thomsen JB, Da Fonseca GAB, Olivieri S (1998) Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conservation Biology*, **12**, 516-520.
- Morán-Ordóñez A, Suárez-Seoane S, Elith J, Calvo L, de Luis E (2012) Satellite surface reflectance improves habitat distribution mapping: a case study on heath and shrub formations in the Cantabrian Mountains (NW Spain). *Diversity and Distributions*, **18**, 588-602.
- Munguía M, Rahbek C, Diniz-Filho JAF, Araújo MB Equilibrium of Global Amphibian Species Distributions with Climate. *PLoS ONE*, e34420.
- Muñoz J, Felicísimo ÁM (2004) Comparison of statistical methods commonly used in predictive modelling. *Journal of Vegetation Science*, **15**, 285-292.
- Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GA, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853-858.

- Nakazawa Y, Williams R, Peterson AT, Mead P, Staples E, Gage KL (2007) Climate change effects on plague and tularemia in the United States. *Vector-borne and zoonotic diseases*, **7**, 529-540.
- Nenzén H, Araújo M (2011) Choice of threshold alters projections of species range shifts under climate change. *Ecological Modelling*, **222**, 3346-3354.
- Ninyerola M, Serra-Díaz JM i Lloret F. 2010. Topo-climatic Suitability Atlas of Woody Plants. Map server. Universitat Autònoma de Barcelona. URL: <http://www.opengis.uab.cat/IdoneitatPI/index.html>
- Ogden N, Maarouf A, Barker I et al. (2006) Climate change and the potential for range expansion of the Lyme disease vector *Ixodes scapularis* in Canada. *International journal for parasitology*, **36**, 63-70.
- Osborne PE, Leitão PJ (2009) Effects of species and habitat positional errors on the performance and interpretation of species distribution models. *Diversity and Distributions*, **15**, 671-681.
- Papes M, Peterson AT, Powell GVN (2012) Vegetation dynamics and avian seasonal migration: clues from remotely sensed vegetation indices and ecological niche modelling. *Journal of Biogeography*, **39**, 652-664.
- Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361-371.
- Peterson AT (2009) Shifting suitability for malaria vectors across Africa with warming climates. *BMC infectious diseases*, **9**, 59.
- Peterson AT (2005) Predicting potential geographic distributions of invading species. *Current science-bangalore*, **89**, 9.
- Peterson AT, Shaw J (2003) *Lutzomyia* vectors for cutaneous leishmaniasis in Southern Brazil: ecological niche models, predicted geographic distributions, and climate change effects. *International journal for parasitology*, **33**, 919-931.
- Peterson AT, Soberon J, Pearson RG, Anderson RP, Martinez-Meyer E, Nakamura M, Araujo MB (2011) *Ecological niches and geographic distributions*. Princeton Univ Pr, .

- Peterson AT (2011) Ecological niche conservatism: a time-structured review of evidence. *Journal of Biogeography*, **38**, 817-827.
- Peterson AT, Papeş M, Soberón J (2008) Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, **213**, 63-72.
- Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259.
- Phillips SJ, Dudík M, Elith J, Graham CH, 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.
- Pulliam HR (2000) On the relationship between niche and distribution. *Ecology Letters*, **3**, 349-361.
- Rangel T, Diniz-Filho J, Araujo M (2009) BIOENSEMBLES 1.0. Software for Computer Intensive Ensemble Forecasting of Species Distributions Under Climate Change. *Goiás, Madrid, Évora*.
- Raxworthy CJ, Martinez-Meyer E, Horning N, Nussbaum RA, Schneider GE, Ortega-Huerta MA, Peterson AT (2003) Predicting distributions of known and unknown reptile species in Madagascar.
- Reed KD, Meece JK, Archer JR, Peterson AT (2008) Ecologic niche modeling of *Blastomyces dermatitidis* in Wisconsin. *PLoS One*, **3**, e2034.
- Regan HM, Syphard AD, Franklin J, Swab RM, Markovchick L, Flint AL, Flint LE, Zedler PH (2011) Evaluation of assisted colonization strategies under global change for a rare, fire-dependent plant. *Global Change Biology*, **18**, 936-947.
- Rota CT, Fletcher RJ, Evans JM, Hutto RL (2011) Does accounting for imperfect detection improve species distribution models? *Ecography*, **34**, 659-670.
- Roura-Pascual N, Suarez AV, Gómez C, Pons P, Touyama Y, Wild AL, Peterson AT (2004) Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **271**, 2527-2535.

- Roura-Pascual N, Suarez AV, McNyset K et al. (2006) Niche differentiation and fine-scale projections for Argentine ants based on remotely sensed data. *Ecological Applications*, **16**, 1832-1841.
- Roura-Pascual N, Hui C, Ikeda T et al. (2011) Relative roles of climatic suitability and anthropogenic influence in determining the pattern of spread in a global invader. *Proceedings of the National Academy of Sciences*, **108**, 220-225.
- Royle JA, Nichols JD, Kéry M (2005) Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, **110**, 353-359.
- Saupe EE, Barve V, Myers CE et al. (2012) Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling*, **237–238**, 11-22.
- Segurado P, Araujo MB (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555-1568.
- Serra-Diaz JM, Ninyerola M, Lloret F (2011) Coexistence of *Abies alba* (Mill.)–*Fagus sylvatica* (L.) and climate change impact in the Iberian Peninsula: A climatic-niche perspective approach. *Flora-Morphology, Distribution, Functional Ecology of Plants*, **207**, 10-18.
- Sharpe PJA (1990) Forest modeling approaches: compromises between generality and precision. Eds. Dixon, R.K., Meldahl, R.S., Ruark, G.A., Warren, W.G., *Process Modeling of Forest Growth Responses to Environmental Stress*. Timber Press, Portland, OR, pp. 180–190.
- Sillero N (2011) What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, **222**, 1343-1346.
- Sinclair SJ, White MD, Newell GR (2010) How useful are species distribution models for managing biodiversity under future climates. *Ecology and Society*, **15**, 8.
- Soberón J (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115-1123.
- Soberón J, Nakamura M (2009) Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, **106**, 19644-19650.
- Somodi I, Molnár Z, Ewald J (2012) Towards a more transparent use of the potential natural vegetation concept—an answer to Chiarucci et al. *Journal of Vegetation Science*, **23**, 590-595.

- Stanton, J. C., Pearson, R. G., Horning, N., Ersts, P. and Reşit Akçakaya, H. (2012), Combining static and dynamic variables in species distribution models under climate change. *Methods in Ecology and Evolution*, **3**, 349–357.
- Sydes NW, Osborne PE (2011) Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography*, **20**, 904-914.
- Syphard AD, Franklin J (2010) Species traits affect the performance of species distribution models for plants in southern California. *Journal of Vegetation Science*, **21**, 177-189.
- Syphard AD, Franklin J (2009) Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. *Ecography*, **32**, 907-918.
- Thomas CD, Cameron A, Green RE et al. (2004) Extinction risk from climate change. *Nature*, **427**, 145-148.
- Thompson M, Zucchini W (1989) On the statistical analysis of ROC curves. *Statistics in medicine*, **8**, 1277-1290.
- Thompson GD, Robertson MP, Webber BL, Richardson DM, Le Roux JJ, Wilson JR (2011) Predicting the subspecific identity of invasive species using distribution models: *Acacia saligna* as an example. *Diversity and Distributions*, **17**, 1001-1014.
- Thuiller W, Araujo MB, Lavorel S (2004a) Do we need land - cover data to model species distributions in Europe? *Journal of Biogeography*, **31**, 353-361.
- Thuiller W, Araújo MB, Pearson RG, Whittaker RJ, Brotons L, Lavorel S (2004b) Biodiversity conservation: uncertainty in predictions of extinction risk. *Nature*, **430**.
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD : a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369-373.
- Thuiller W, Lavorel S, Araújo MB, Sykes MT, Prentice IC (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 8245-8250.
- Triviño M, Thuiller W, Cabeza M, Hickler T, Araújo MB (2011) The Contribution of Vegetation and Landscape Configuration for Predicting Environmental Change Impacts on Iberian Birds. *PloS one*, **6**, e29373.

- Udvardy MDF (1959) Notes on the ecological concepts of habitat, biotope and niche. *Ecology*, **40**, 725–728.
- US Global Research Act -USGCRP (1990) P.L. 101-606
- Václavík T, Meentemeyer RK (2012) Equilibrium or not? Modelling potential distribution of invasive species in different stages of invasion. *Diversity and Distributions*, **18**, 73-83.
- VanDerWal J, Shoo LP, Johnson CN, Williams SE (2009a) Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist*, **174**, 282-291.
- VanDerWal J, Shoo LP, Graham C, Williams SE (2009b) Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, **220**, 589-594.
- Veloz SD (2009) Spatially autocorrelated sampling falsely inflates measures of accuracy for presence - only niche models. *Journal of Biogeography*, **36**, 2290-2299.
- Veloz SD, Williams JW, Blois JL, 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.
- Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335-342.
- Wiens JJ, Ackerly DD, Allen AP et al. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310-1324.
- Williams JN, Seo C, Thorne J, Nelson JK, Erwin S, O'Brien JM, Schwartz MW (2009) Using species distribution models to predict new occurrences for rare plants. *Diversity and Distributions*, **15**, 565-576.
- Williams JW, Jackson ST (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, **5**, 475-482.
- Wissel C (1992) Aims and limits of ecological modeling exemplified by island theory. *Ecological Modelling*, **63**, 1–12.
- Wisz MS, Hijmans R, Li J, Peterson AT, Graham CH, Guisan A (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763-773.

- Yeshiwondim AK, Gopal S, Hailemariam AT, Dengela DO, Patel HP (2009) International Journal of Health Geographics. *International journal of health geographics*, **8**, 5.
- Zimmermann N, Edwards Jr T, Moisen G, Frescino T, Blackard J (2007) Remote sensing - based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology*, **44**, 1057-1067.
- Zurell D, Elith J, Schröder B (2012) Predicting to new environments: tools for visualizing model behaviour and impacts on mapped distributions. *Diversity and Distributions*, **18**, 628-634

2. FIRST STEPS

2.1 The topoclimatic Atlas of the Iberian Peninsula

Summary: Topo-climatic suitability maps of the main woody plant species in raster format in a 200 m spatial resolution. Data used has been obtained from the Digital Climatic Atlas of the Iberian Peninsula (DCAIP; see Figure 2.1) and the third National Forest Inventory (NFI3). Using a GLM regression as a classifier (General Linear Model), climatic information has been combined with species locations using GIS tools (MiraMon) to map its suitability. OGC standards (Open geographic consortium) have been used to its Internet publication ensuring interoperability with other servers. Results shown are congruent with different modeling techniques used for the same study area (Benito-Garzon et al. 2008).

Authorship: Ninyerola M, Serra-Díaz JM & Lloret F. 2010. Topo-climatic Suitability Atlas of Woody Plants. Map server. Universitat Autònoma de Barcelona.

URL: <http://www.opengis.uab.cat/ldoneitatPI/index.html>

Technical support: Alex Franquesa, Núria Julià, Meritxell Batalla.

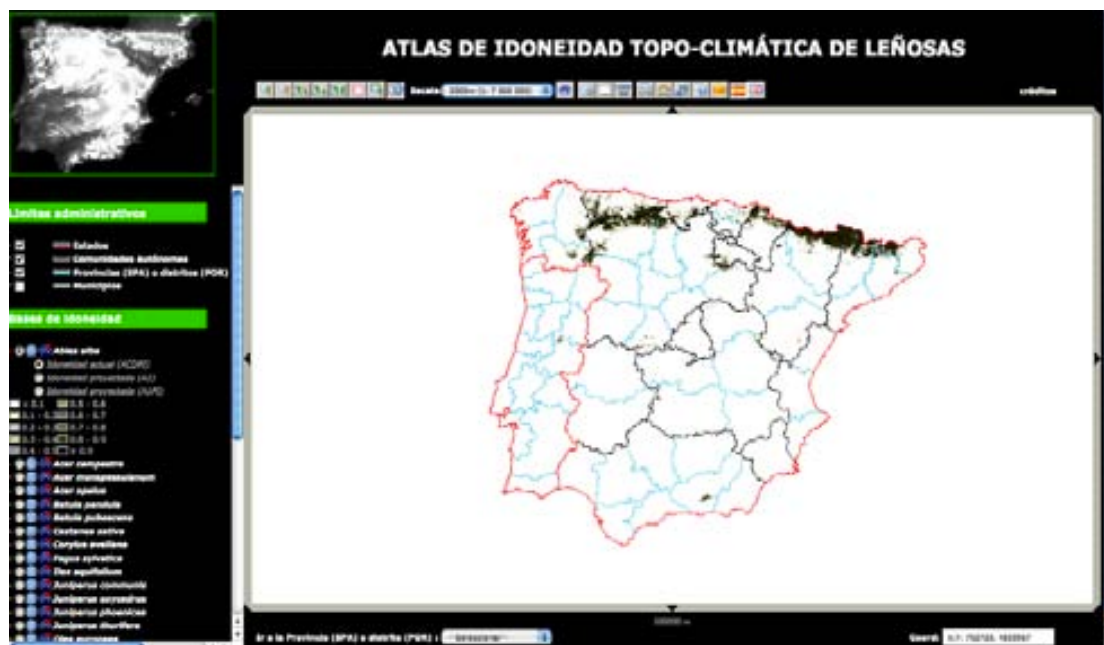


Figure 2.1 Web display of the Atlas of Topo-climatic suitability.

2.2 Material and methods

General Linear Models (GLM) have been used in order to model topo-climatic suitability of woody plants using presence and absence data from the plots of the National Forest Inventory (MMA 1994-2004). Although there are a wide range of classification methods (Araújo & New 2007; Thuiller et al. 2009), GLMs constitute one of the most used techniques for modeling species distribution (Guisan & Zimmermann 2000; Randin et al. 2009) given its ability to model realistically ecological relationships (Austin, 2002; Elith et al., 2006).

Presence-Absence data of species.

Plots of the third National Forest Inventory have been used to identify locations where species is present or absent. For each species, presence or absence has been used as the dependent variable.

Data on absences may influence model accuracy (Lobo et al., 2008) and the kind of distribution that we want to model (realized or potential ecological niche) (Kearny, 2006). Given that the number of absences is larger than the number of presences for a target species, 250 datasets have been build where the number of presences and the number of absences are balanced, hence obtaining a robust model not dependent on prevalence (but see Chefaoui & Lobo, 2008).

Species absence does not per se ensure its unsuitability because many other factors may be affecting its absence, besides climate and/or topography. Therefore it may be deduced that we are using pseudo-absences (we do not know whether the species is not present due to climatic factor or due to other factors). In order to diminish such an effect, we have selected as absences those plots where the species is not present and at a minimum distance of 5 km of a presence plot.

Selection of topographic and climatic variables

Climatic variables were derived from the Digital Climatic Atlas of the Iberian Peninsula (DCAIP) (Ninyerola et al., 2007a,b). This cartography consists of 65 monthly maps of mean air temperature (minimum, mean and maximum values), precipitation and solar radiation generated through spatial

interpolation techniques from c.a. 2000 meteorological stations (for precipitation) and c.a. 1000 meteorological stations (for temperature).

Furthermore, monthly and annual water balances have been calculated as the difference between precipitation and potential evapotranspiration (using Thornthwaite) without considering runoff.

Topographic variables, for instance slope, curvature, solar radiation and friction (cost-distance to cost), have been calculated using a Digital Elevation Model (DEM) of 200m spatial resolution. This sort of variables has been shown to be relevant for forest distribution (del Barrio et al. 1997; Pfeffer et al., 2003; Bailey, 2004).

A correlation analysis has been undertaken in order to reduce colinearity in models and variables presenting more than 0.70 correlation have been suppressed from the model. In order to choose which variable among the correlated variables should be used, we decided to introduce the most integrative variable. For example, we systematically maintained water balance (highly correlated with precipitation and temperature).

Evaluation

These models have realized with the 80% of plots keeping the 20% for cross validation. Due to use of 250 subsets of presence-absence data, evaluation results express the mean of the 250 models for each species.

The calculated index for evaluation is the area under the ROC curve (AUC) (Receiver Operating Characteristics). Such an index varies between 0 and 1. In general, AUC values over 0.80 mean that the model used is a good classifier and therefore, an accurate prediction of suitability.

Future scenarios: General Circulation Models (GCM)

Socio-economic scenarios A1FI and A2 from HadCM3 simulation have been used. These simulation is linked to a coupled atmosphere-ocean circulation model (AOGCM) developed at the Hadley Centre – UK (Gordon et al., 2000; Pope et al., 2000; IPCC, 2007).

These models predict monthly and annual temperature and precipitation for the 2050-2080 period in a 4 km spatial resolution. Future regionalized climate was obtained using an approximation based on differences between the past climate (CRU) and the climate projection from the HADCM3 model using the A1 and A2 storyline, thus combining the predictive information of the GCM with the topoclimatic data provided by ground stations.

2.3 References and software

- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42-47
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**,101-118.
- Bailey R (2004) Identifying Ecoregion Boundaries. *Environmental management*, **34**, S14-S26.
- Benito Garzón M, Sánchez de Dios R, Sainz Ollero H (2008) Effects of climate change on the distributions of Iberian forests. *Applied Vegetation Science* ,**11** ,169-178.
- Chefaoui RM, Lobo JM (2008) Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling*, **210**, 478-486.
- Elith J, Graham CH, Anderson RP, et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151.
- del Barrio G, Alvera B, Puigdefabregas J, Diez C (1997) Response of high mountain landscape to topographic variables: Central Pyrenees. *Landscape Ecology*, **12**, 95-115.
- Gordon C, Cooper C, Senior CA, et al. (2000) The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, **16**, 147-168.
- Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147-186.
- 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 [Solomon, S, D Qin, M Manning, Z

- Chen, M Marquis, KB Jones, PD, New M, Parker DE, Martin S, Rigor IG (1999) Surface air temperature and its variations over the last 150 years. *Review of Geophysics*, **37**,173–199.
- Kearny M (2006) Habitat, environment and niche: what are we modelling? *Oikos*, **115**, 186-191.
- Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145-151.
- MMA, Organismo Autónomo de Parques Nacionales. 1999-2004. Publicaciones del Tercer Inventario Forestal Nacional. Madrid.
- Ninyerola M, Pons X, Roure JM (2007a) Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. *International Journal of Climatology*, **27**, 1231-1242.
- Ninyerola M, Pons X, Roure JM (2007b) Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. *Theoretical and Applied Climatology*, **89**, 195-209.
- Pfeffer K, Pebesma E, Burrough P (2003) Mapping alpine vegetation using vegetation observations and topographic attributes. *Landscape Ecology*, **18**, 759-776.
- Pope V, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parameterizations in the Hadley Centre climate model: HadAM3. *Climate Dynamics*, **16**, 123-146.
- Randin CF, Engler R, Normand S et al. (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, **15**, 1557-1569.
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD - a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369-373.

SOFTWARE USED

- Pons X (2004) MiraMon. Geographical information system and remote sensing software. Versión 5. Centre de Recerca Ecològica i Aplicacions Forestals (CREAF). ISBN: 84-931323-4-9. <http://www.creaf.uab.es/miramom>

R Development Core Team (2009). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

- <http://www.R-project.org>

2.4 Acknowledgements and participants

Guadalupe Barea (BB Disseny Digital) for text and web design.

This server has been developed in the framework of the Spanish Ministry of Science and Technology, National Plan R+D+I, Project Climate Change and woody plant vulnerability at different scales: regional distribution and local population dynamics” (CGL2006-01293/BOS) and the project Consolider Montes (CSD2008-00040).

We would like to especially thank those who have enabled us to access to fundamental data.

National Forest Inventory

We would like to Roberto Vallejo and their team in the Nature Data Bank Area (Dirección General para la Biodiversidad. Ministerio de Medio Ambiente).

We also express here our gratitude to Jordi Vayreda and Juanjo Ibañez (CREAF).

Current climate scenario: <http://www.opengis.uab.es/wms/iberia/index.htm>

Future climate scenarios: Santi Sabaté (UB and CREAM).

ALARM (Assessing Large-scale environmental Risks for biodiversity with tested Methods: and ATEAM (Advanced Terrestrial Ecosystem Analysis and Modelling: EVK2-2000-00075) project, from the EU Fifth Framework for Energy, Environment and Sustainable Development.

3. ON CURRENT DISTRIBUTIONS

STUDY: ‘Geographical patterns of congruence and incongruence between niche-based and eco-physiological models’¹

3.1 Introduction

Species distributions are driven by complex interactions between biotic and abiotic factors. Models constitute useful tools for describing patterns of species distribution, while also contributing to the identification of driving factors, the prediction of future species distribution and the weighing up of decisions about conservation in complex multidisciplinary scenarios (Thuiller et al., 2005a; Falk & Mellert, 2011). Species-specific models can be classified into two distinct categories: empirical niche-based models and process-based models (Kearney, 2006). These two approaches represent contrasting methodologies, and when applied together they can be used to provide insight into changing conditions (Morin & Thuille 2008; Keenan et al., 2011a; Dormann et al., 2012).

Niche-based models (NBM, also known as species distribution models, bioclimatic envelopes, correlative envelope models) comprise a family of empirical statistical modelling approaches and data-mining techniques that correlate environmental variables with the presence and/or absence of a species in order to determine its distribution. These models profile the bioclimatic envelope of the species via a number of different statistical techniques and then project these conditions onto the geographical space (Franklin, 2010). This correlative approach has extensively been used to address many questions in ecology and conservation: identification of hotspots, understanding of speciation (Graham et al., 2004), assessment of extinctions (Thomas et al., 2004) and alien species invasions (Thuiller et al., 2005b), and projection of the effects of climate change on species

¹ This study was led by JM Serra-Diaz, who performed the correlative ecological niche models and first draft of the study. Dr.Keenan performed the analysis and described results for physiological process-based model. Conclusions are drawn from the interaction of the two researchers with great input from other co-authors.

distribution (Keenan et al., 2011a). Perhaps the most fundamental criticism is the need to assume the equilibrium or pseudo-equilibrium of species distribution with climate (Austin, 2002), even though many taxa have proved to be at different degrees of equilibrium with climate (Araújo & Pearson, 2005; Svenning & Skov, 2004).

Alternatively, process-based models (PBM, also biophysical models) represent a modelling approach focused on species performance in a given environment, providing insights on the mechanisms that directly or indirectly shape species distribution (e.g. growth, dispersal, reproduction, etc.). They use detailed information on the target species or group to model a certain ecological or physiological process at the species level or at the functional level. For terrestrial vegetation, PBMs have been developed to describe, for instance, fitness (Chaine & Beaubien, 2001) or growth (Rickebusch et al., 2008). They can integrate different processes on a broader scale, such as landscape (Scheller & Mladenoff, 2007), gap models (Bugmann, 2001), global vegetation dynamics (Sitch et al., 2003) and the techniques used may range from empirical relationships to physical laws.

Interestingly, this family of models has barely been adopted for predicting species distribution (but see Jeltsch et al. (2008); Buckley (2008) and Kearney et al., (2008)), as their widespread use may be hampered by the computational power and data required, thereby restricting their application to well-known taxa (Ainsworth et al., 2008). Furthermore, PBMs are subject to high parameter dimensionality, with associated complications in quantifying uncertainty in model projections (Raupach et al., 2005; Keenan et al., 2011b).

Comparisons between the two modelling approaches (NBMs vs. PBMs) are rare, and they have generally focused on testing model predictions under climate change scenarios. For instance, Keenan et al. (2011a), Morin & Thuiller (2009) and Rickebusch et al. (2008) have shown that NBMs potentially overestimate the negative effect of climate change on current forest stands, due to CO₂ fertilization and plant water availability. Few studies have experimentally addressed the link between physiology and NBM-generated values of suitability (but see Austin et al. (2009)). On broad scales, and depending on the goal, it has been argued that general climatic variables prove sufficient to evaluate NBM models (Araújo & Guisan, 2006), but it has also been recognised that the

identification of more relevant predictors of species distributions requires a greater understanding of the interplay between physiology, climate variables and scale (Elith & Leathwick, 2009; Austin & Van Niel, 2011). In fact, an already validated PBM allows us to test such hypothesis projected in current environments where the species is not present at the moment, thus representing a *controlled experiment* compared to the complex analysis of inventory data, where a lot of signals (management, climate, etc.) could lead to confusion.

In the present study we compare predictions from NBM (habitat suitability) and PBM (growth) for three common Mediterranean tree species (*Pinus halepensis*; *Quercus ilex*; *Pinus sylvestris*) in the Iberian Peninsula (Figure 3.1a), through a range of environments (arid to mesic to mountain conditions respectively). Specifically, our aims are to (1) assess whether habitat suitability and growth show similar responses across the three species and (2) map regions of agreement and disagreement between the two approaches. We hypothesise that, if species distribution is in equilibrium with climate and in absence of biotic interactions, a robust positive relationship between NBM measures of habitat suitability (occurrence) and PBM measures of growth would be found (Figure 3.1b).

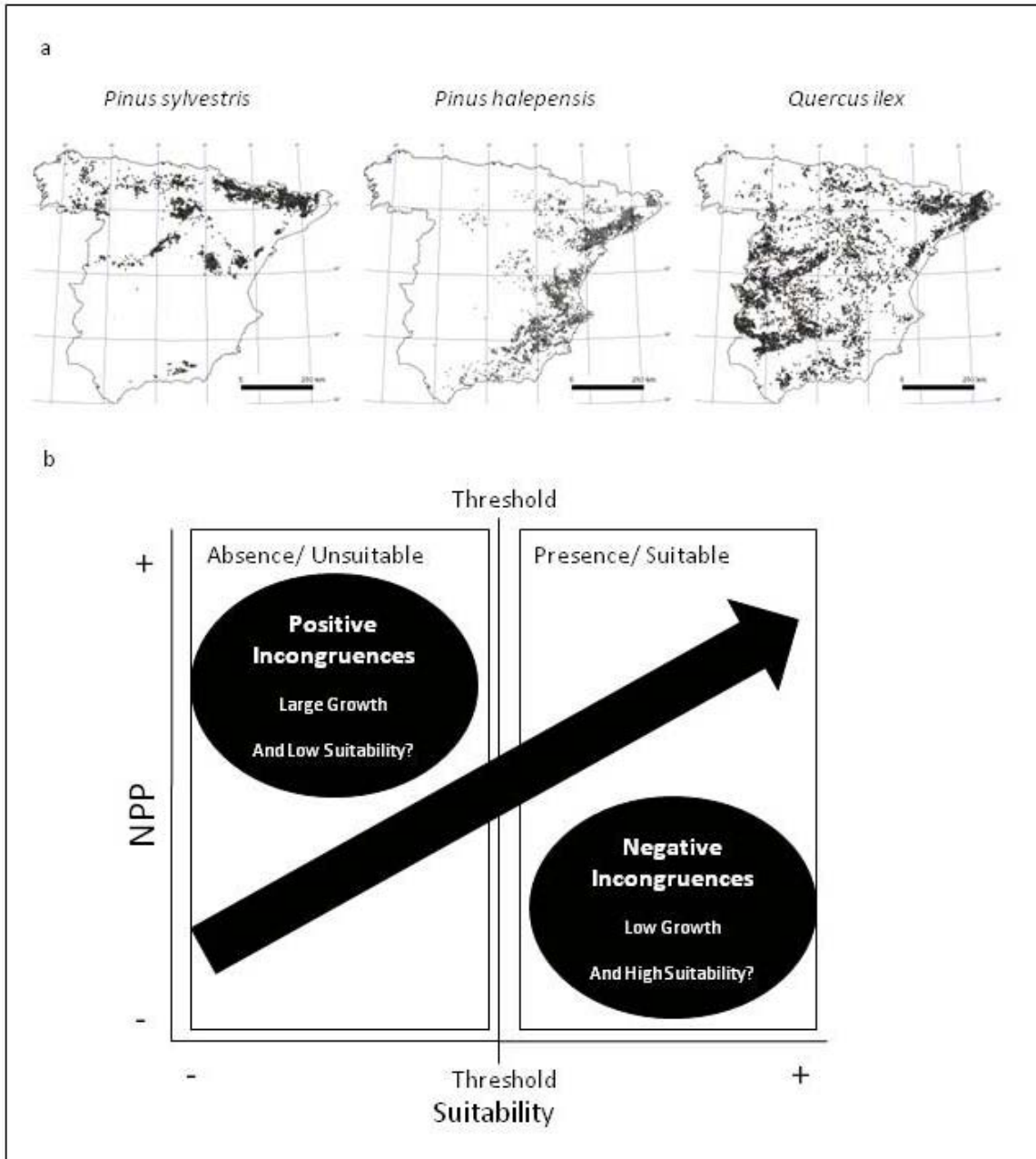


Figure 3.1 (a) Current distribution in Spain of *Pinus sylvestris*, *Pinus halepensis* and *Quercus ilex* according to the Spanish Forest Inventory (SFI) (b) Black arrow indicates expected relationship between growth and site suitability if species distribution is broadly in equilibrium with climate. Two types of incongruence are expected: positive incongruence (I^+), when high growth performance corresponds to low suitability estimates, and negative incongruence (I^-), when low growth corresponds to high suitability.

3.2 Material and methods

Niche-based models (NBM)

In order to attain robust results and quantify uncertainty resulting from differences between NBMs (Elith & Graham, 2009), we used a group of seven different NBMs (Araújo & New, 2007) with the BIOMOD platform (Thuiller et al., 2009) implemented in R statistical software (R Development Core Team 2011). All these models (see Table 3.1) require data on species presence and absence to be related to environmental variables. Presence and absence records were extracted from the third Spanish Forest Inventory (SFI; MinisterioAPyA, 2007), which surveyed forest species at a 1-km regular grid spatial resolution resulting in 10 784, 18 268 and 10 202 presence plots for *Pinus halepensis*, *Quercus ilex* and *Pinus sylvestris*, respectively, from the total of 70 855 plots in the Spanish Forest Inventory. We applied a random selection of absences from the forest inventory equal to two times the number of presences, allowing for a large number coverage of the environmental space (Barbet-Massin et al., 2012). The BIOMOD-platform automatically weights presences and absences so that the weighted sum of records maintains a prevalence of 0.5, hence contributing equally to the model construction.

The climatic variables used in the NBMs are directly related to growth: mean air temperature, annual precipitation and solar radiation. Climate data were extracted from the digital climatic atlas of the Iberian Peninsula (http://opengis.uab.es/wms/iberia/en_index.htm; Ninyerola et al., 2000; Ninyerola et al., 2007a; Ninyerola et al., 2007b). It retrieves monthly mean climatic data at a 200-m spatial resolution. Although other variables have been used to derive NBMs on a regional scale, i.e. topography (Randin et al., 2009; Austin & Van Niel, 2011), we preferred using the variables most similar to the PBM. In addition, previous analysis using other set of variables retrieved similar results (Ninyerola et al., 2010; <http://www.opengis.uab.es/wms/IdoneitatPI/index.htm>). The most important variables affecting outputs were precipitation and mean annual temperature, for *Pinus sylvestris* and *Quercus ilex* respectively. Both variables had a very similar influence on *Pinus halepensis* NBM (see Appendix 2.1 in Supporting Information).

MODEL	DESCRIPTION	REFERENCE
RF (random forest)	A machine-learning method – a combination of tree predictors in which each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest.	Breiman (2001)
CTA (classification tree analysis)	A classification method – a 50-fold cross-validation to select the best trade-off between the number of leaves of the tree and the explained deviance.	Breiman et al. (1984)
MARS (multivariate adaptive regression splines)	A non-parametric regression method, mixing CTA and GAM.	Friedman (1991)
GLM (generalised linear model)	A regression method, with polynomial terms for which a stepwise procedure is used to select the most significant variables.	McCullagh and Nelder (1989)
GAM (generalised additive model)	A regression method, with 4 degrees of freedom and a stepwise procedure to select the most parsimonious model.	Hastie&Tibshirani (1990)
GBM (Generalised Boosting Models)	A method that fits a large tree of simple models, together aimed at giving a more robust estimate of the response. Based on Boosted Regression Tree algorithm.	Friedman (2001)
ANN (artificial neural networks)	A machine-learning method, with the mean of three runs used to provide predictions and projections.	Ripley (1996)

Table 3.1 NBM used in this study. See the BIOMOD manual (Thuiller et al, 2009).

Model calibration was undertaken with 70% of the presence-absence stands and the remaining 30% were used for validation. The number of runs was set to three, with two repetitions of each run using different absence combinations; there were thus a total of six suitability outputs for each. The models project a suitability index that ranges from 0 to 1 000 for each location under consideration. In order to discriminate suitable from unsuitable conditions, a threshold needs to be set and there are a number of methods for this, based on accuracy measurements (Nenzén & Araújo, 2011). In the present study we set the threshold by maximizing the True Skill Statistic (TSS) metric, as this is not influenced by prevalence (Allouche et al., 2006), although the area under the ROC curve and Kappa (other accuracy measures) were also reported (see Appendix 2.1). They overall retrieve good accuracy results showing that the NBMs are able to model current distribution.

Robust suitability measures were obtained from the ensemble of the seven calculated models. This procedure can be undertaken by using a wide range of approaches (Araújo & New, 2007). In the present study we employed the TSS accuracy measure to estimate each model’s contribution to the ensemble. We first ranked models by TSS accuracy and then applied a down-weighting factor of

1.6, which implies that the weight of each model will be 1.6 times smaller than the previous one (see Thuiller et al., (2009) for further details). Such values represents a compromise between null weighting (decay = 1) and overweighting (decay=2).

Process-based models (PBM)

GOTILWA+ (Growth Of Trees Is Limited by Water; Gracia et al., 1999; Sabate et al., 2002; Keenan et al., 2011a; <http://www.creaf.uab.es/GOTILWA+>) is a process-based terrestrial biogeochemical model of forest growth. Developed for the Mediterranean region, it is has been used to explore how forests are influenced by water stress, tree stand structure, management techniques, soil properties, and climate (including CO₂) change.

GOTILWA+ does not predict the distribution of a species, but simulates tree growth, and the associated carbon and water fluxes, for different tree species in different environments. Growth, or modeled net primary production (NPP hereafter), can be interpreted as reflecting a site-species specific ecophysiological value comparable to estimates of suitability generated by NBMs (Keenan et al., 2011a). The model treats mono-specific stands, which can be even or uneven-aged. Individual trees are aggregated into 50 dbh (diameter at breast height) classes and calculations are performed separately for each class. Hourly ecosystem carbon and water fluxes are estimated in dependence on meteorological forcing and ecosystem state.

The GOTILWA+ model includes a two-leaf canopy photosynthetic sub-model (Wang & Leuning, 1998; Dai et al., 2004) for the C₃ photosynthetic pathway. The canopy is divided into sunlit and shaded leaves, with the amount of intercepted diffuse and direct radiation depending on the time of the day, season, and the area of leaf exposed to the sun (Campbell, 1986). Foliage net assimilation rates are calculated by using the Farquhar et al. (1980) photosynthesis model, with dependencies on intercepted quantum flux density, species-specific photosynthetic capacities, leaf temperature, and leaf intercellular CO₂ concentration. Stomatal conductance is calculated by using the Leuning et al. model (Leuning et al., 1995), which is a version of the original Ball et al. (1987) model. Net photosynthesis is scaled from the leaf to the canopy through the canopy microclimate model to give canopy bulk gross primary production (GPP). NPP is calculated as the balance of GPP

minus autotrophic respiration components and is defined as $NPP = GPP - (R_f + R_w + R_r)$, where R_f is night respiration rate per ground unit area, R_w is respiration of non-leaf aerial plant tissues, R_r is respiration of root tissues. Model parameters were set to species-specific values (as in Gracia et al., 1999; Keenan et al., 2009a,b, 2011).

Water stress affects the photosynthesis–conductance coupling by directly reducing the photosynthetic potential through a non-linear relationship with soil water content (Keenan et al., 2009a). Phenology is temperature-dependent and accounted for in an updated version of the Pelkonen & Hari (1980) approach for calculating the seasonal variations in photosynthetic potential. GOTILWA+ has been validated and widely applied both in the Mediterranean region and the rest of Europe (see Morales et al., 2005; Keenan et al., 2009a for validation exercises and Keenan et al., 2009b, c, 2011 for example applications).

Study design

We employed a virtual data approach by creating virtual mono-specific forests (VF) for each species (*Pinus halepensis*, *Pinus sylvestris* and *Quercus ilex*). VF structure traits (e.g., tree density, diameter distributions, soil depth) were set to the mean of species-specific forest structural characteristics derived from the third Spanish forest inventory stands. Simulations of VFs' growth were performed across the whole gradient of suitability predicted by the NBMs. First, 3 000 virtual forests were randomly assigned to each of the two different regions of the suitability index determined by the threshold criteria (maximum TSS): not suitable (below threshold) and suitable (above threshold) conditions. Second, we simulated forest growth using GOTILWA+ at these locations.

We explored the relationship between NPP and suitability using robust statistics from *robustbase* package (Rousseeuw et al., 2011) using R statistics software (R Development core team 2011). Robust regression has been widely applied, especially when assumptions of an ordinary least squared are violated, or to identify outliers in the regression (Maronna & Yohai, 2003).

We applied a data-driven approach to quantify the shape of the NPP-suitability relationship using different polynomials. In order to avoid over-fitting, a step-wise backwards procedure was used to find a trade-off between best fit and minimum degree of the polynomial. As a measure of the fit of these regressions, we calculated the R^2 for robust regression (Renaud & Victoria-Feser, 2010). Starting with an arbitrarily high polynomial degree, at each step, every R^2 of the polynomial was compared to the R^2 of the polynomial with one degree less. This process chooses the lowest polynomial degree, unless the difference between two steps is less than 0.05 in R^2 . *Quercus ilex* and *Pinus sylvestris* were best fit, using a second-degree polynomial, whereas *Pinus halepensis* required a fourth-degree.

In order to classify each VF as congruent or incongruent we performed a cluster analysis, using as a distance measure the absolute value of residuals to the fitted robust function. We used k-means clustering by applying the Hartigan & Wong algorithm (1979). We set $k=2$ in order to find two groups of VF: those congruent with the regression and those identified as outliers (incongruences).

Following the trend outlined in Figure 3.1, we finally identified discrepancies using the following criteria:

- Positive incongruence (I^+): VF in the cluster of outliers, predicted as absences by NBMs and with positive residual values. That is, high values of NPP and low climatic suitability from the NBM.
- Negative incongruence (I^-): VF in the cluster of outliers, predicted as presences by NBMs and with negative residual values. That is, with low values of NPP and high climatic suitability from the NBM.

We additionally analysed two physiological features related to growth that are modelled in GOTILWA+: Leaf area index (LAI) and mean leaf life (MLL). Such variables, besides being coupled with NPP, encapsulate species-specific strategies to cope with stressing factors such as drought, which is predominant in the Mediterranean climate.

3.3 Results

General patterns of congruence and incongruence between growth and suitability values.

Quercus ilex and *Pinus sylvestris* present the expected positive relationship (congruence) between NPP and suitability obtained by PBM and NBM, respectively (Figure 3.2a,b). In contrast, *Pinus halepensis* shows a negative relationship, where NPP is lower at high suitability than NPP at low suitability (Figure 3.2). For the three species studied, the range and variability in NPP values decrease along the axis of the suitability index (Figure 3.2). In other words, suitable climatic conditions tend to render more similarly modelled NPP values.

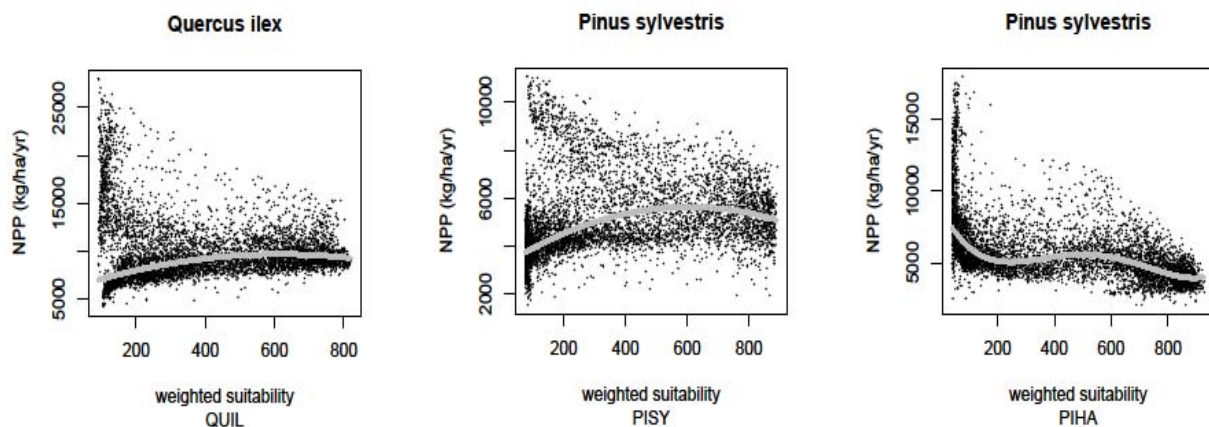


Figure 3.2 Relationship between growth (NPP) and suitability (NBM weighted ensemble) for (a) *Quercus ilex*, (b) *Pinus sylvestris* and (c) *Pinus halepensis*. Line trends represent the fitted robust polynomial regression.

The values of robust R^2 are 0.43 (polynomial degree 4) for *Pinus halepensis*, the species showing no general congruence, 0.35 (polynomial degree 2) for *Pinus sylvestris* and 0.22 (polynomial degree 2) for *Quercus ilex*. The high value for *Pinus halepensis* explained by the degree of the polynomial fit, but it decreases to 0.30 when considering a polynomial degree of 2. Accordingly, we found similar results when analysing the correlation between suitability values of individual NBMs and NPP, this correlation being higher in *Pinus sylvestris* and *Pinus halepensis* than in *Quercus ilex* (Table 3.3).

Incongruence of virtual forests and geographical distribution

We identified incongruence regions only for those species presenting the expected general positive relationship between suitability and NPP (*Quercus ilex* and *Pinus sylvestris*). For *Quercus ilex*, incongruent VFs constitute less than 10% of the total simulated forests (Table 3.2). All discrepancies for this species represent positive incongruences (I^+), meaning that NPP is high despite being in low climatic suitability locations. Geographically, 95% of incongruent areas were located in the Northern Iberian Peninsula, the region receiving the maximum precipitation with subtly lower temperatures than the rest of the area. A few I^+ represent isolated patches in other mountain ranges of the centre of the peninsula (Figure 3b,d).

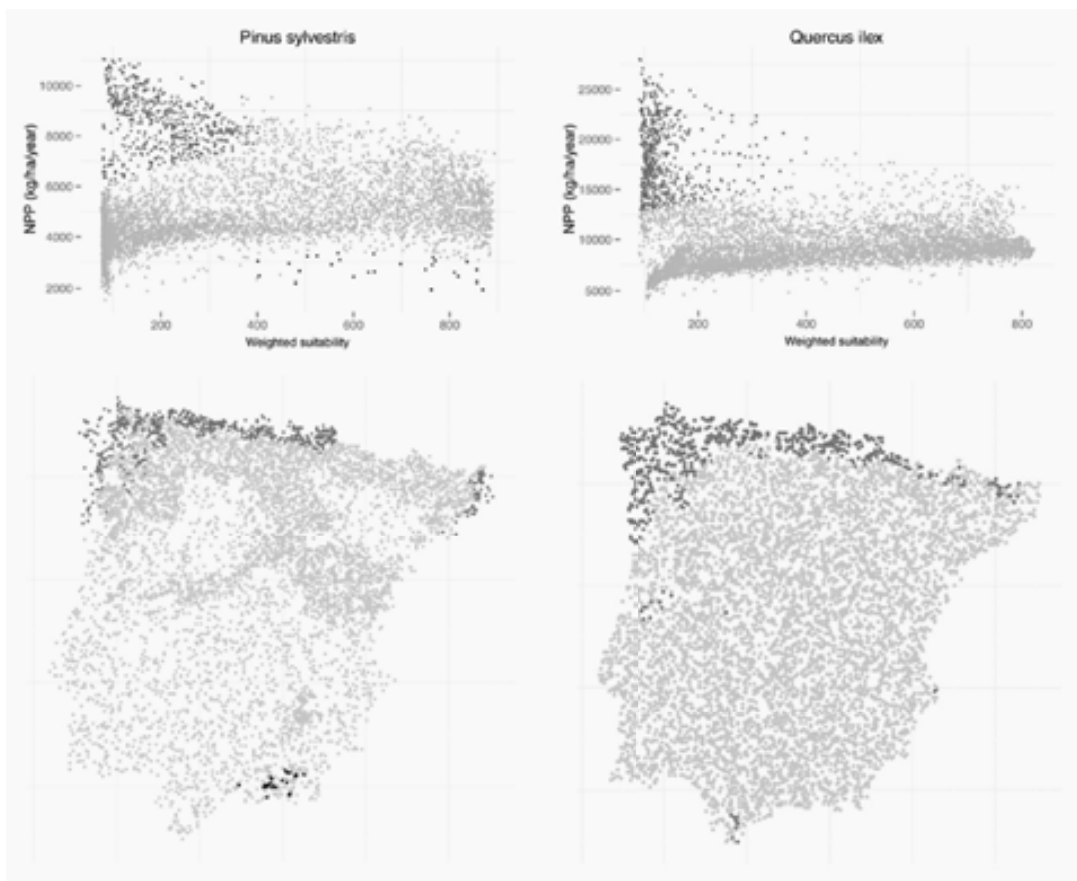


Figure 3.3 Positive and negative incongruence identified for *Pinussylvestris* (a) and *Quercus ilex* (b) and their geographical location (c, d, respectively) along a suitability gradient (weighted ensemble of NBM). Lines indicate density of VFs (virtual forests) of each category.

SPECIES	INCONGRUENCE TYPE	% VFs	T annual (°C)	Pannual (mm)	Rannual (MJ/m ² *day)	Elevation (m)
<i>Quercus ilex</i>	Congruence	90.4%	15.1 (±2.6)	615 (±208)	21.60 (±1.59)	654 (±393)
	Positive incongruence	9.6% (578)	11.4 (±2.1)	1331 (±242)	19.66 (±1.49)	566 (±419)
	Negative incongruence	0%	-	-	-	-
<i>Pinus sylvestris</i>	Congruence	90.6%	12.1 (±2.7)	727 (±285)	20.86 (±1.32)	837 (±420)
	Positive incongruence	9.0% (546)	11.9 (±1.4)	1357 (±277)	19.84 (±1.43)	466 (±280)
	Negative incongruence	0.4% (26)	9.5 (±1.7)	588 (±147)	19.88 (±1.80)	1910 (±354)

Table 3.2 Summary of incongruence values for each species considered and the associated environmental variables. Values correspond to the average of the conditions in the VFs (virtual forests) and values in brackets indicate standard deviation. P annual: annual accumulated precipitation; T annual: mean annual air temperature; R annual: mean potential solar radiation.

For *Pinus sylvestris* both positive (I^+) and negative (I^-) incongruences are observed. However, I^+ (9.0% of the simulated forests) are far more prevalent than I^- (0.4% of the simulated forests) (Table 3.2). The geographical pattern of I^+ is the same as in *Quercus ilex*: a concentration in the Northern Iberian Peninsula (Figure 3.3b,d). These sites are also characterised by higher precipitation and similar temperature than congruent VFs (Table 3.2). I^- are clustered in the South-East range of the Iberian Peninsula (Sierra Nevada) and are characterised by high altitudes with lower precipitation than the majority of VFs of this species.

Physiological responses: Mean leaf life (MLL) and leaf area index (LAI)

For *Quercus ilex*, LAI rapidly increases with suitability and reaches a maximum within the congruence region (Figure 3.4a). However, we observe a cluster of large LAI values at low suitability, which corresponds to I^+ conditions. In these climatic environments, species would react by increasing leaf area and therefore also increasing NPP. It is worth noting that for LAI there is no complete separation between I^+ and the rest of VFs. The relationship between MLL and suitability does not show a clear pattern (Figure 3.4b) as there is great dispersion. However, I^+ generally show relatively higher MLL than the rest of VFs.

For *Pinus sylvestris*, the relationship between LAI and suitability is not straightforward (Figure 3.4c). As in the case of NPP, the variability of LAI values narrows as suitability increases. I^+ conditions are characterised by higher LAI and I^- conditions correspond to lower LAI than the rest of VFs, although in the last case no sharp separation is evident. MLL for this species describes three different trends along a well-defined function: (1) for suitability values below 200, MLL decreases exponentially; (2) for suitability values above 800 it increases exponentially; (3) for suitability values between 200 and 800, MLL subtly increases in a linear fashion. Overall, MLL is in accordance with LAI values and NPP (Figure 3.4d): unsuitable conditions are linked to water stress and unproductive sites where leaves tend to be smaller and remain active for less time on the tree.

Overall, *Pinus halepensis* shows an opposite pattern to that of the other two species. As with NPP, LAI tends to be lower at higher suitability (Figure 3.4e), even though the variability is very large. MLL presents two opposite patterns along the suitability gradient (Figure 3.4f): a positive linear relationship at MLL values below around 3 years, and a negative linear relationship above this value. These two opposite trends merge around suitability values of 600. The negative relationship corresponds to more humid conditions in which the species is in fact not present but water and light are available (North Iberian Peninsula; Appendix 2.2). Under these conditions, this species would tend to maintain leaves, thus resulting in high NPP at low suitability. The positive relationship corresponds to VFs placed in xeric conditions (Southern Iberian Peninsula; Appendix 2.2), where leaves are not maintained.

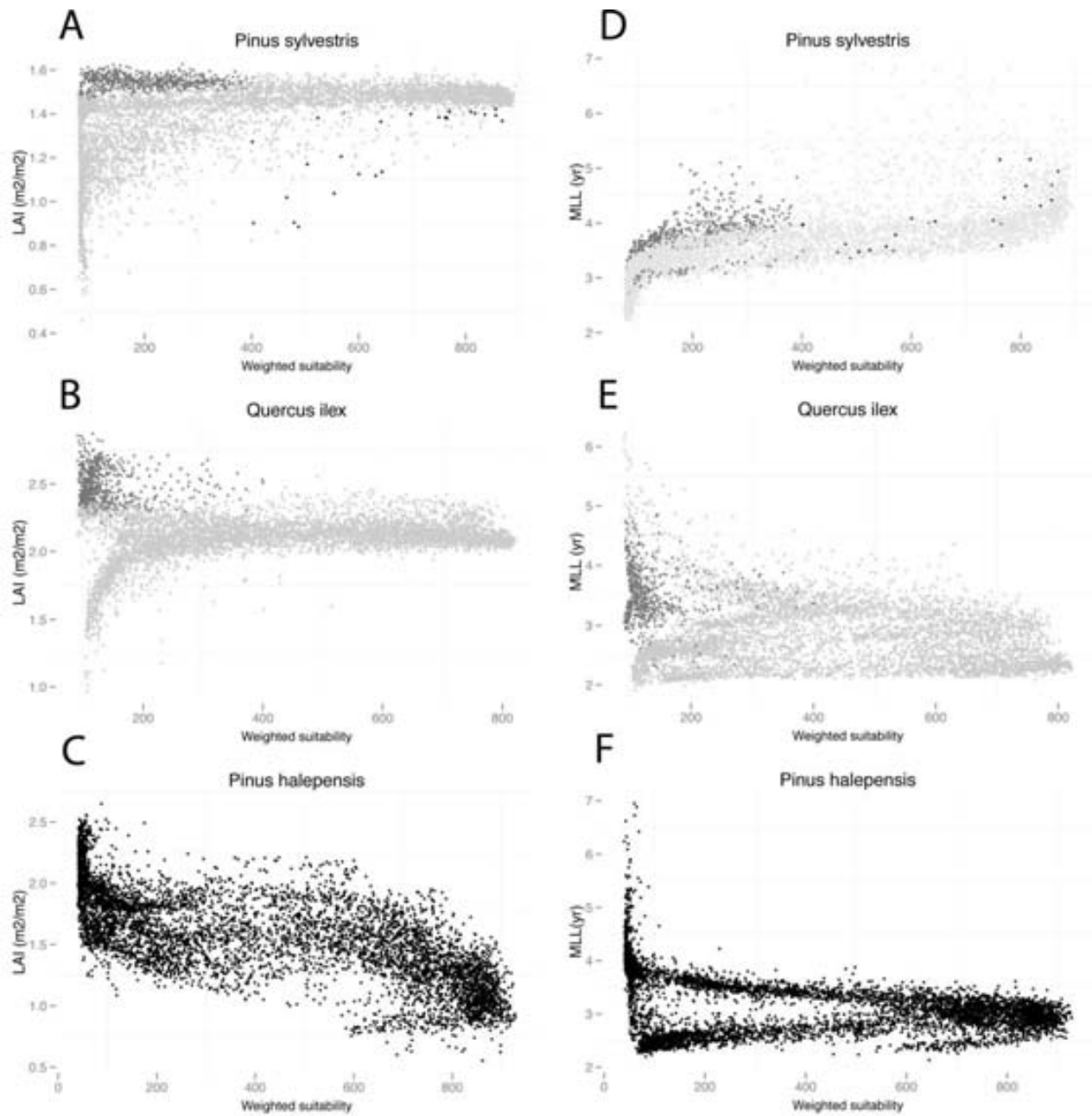


Figure 3.4 Relationship of Leaf Area Index (LAI) and Mean Leaf Life (MLL) with suitability (NBM weighted ensemble), for *Quercus ilex* (a, b), *Pinus sylvestris* (c, d) and *Pinus halepensis* (e, f). Dark grey colours indicate VFs (virtual forests) that show a positive incongruence and light grey indicates VFs that show negative incongruence. Black points correspond to congruence cases. Note that *Pinus halepensis* has no categories due to the overall pattern of incongruence.

Differences between Niche-based model techniques and the role of the threshold value.

We computed correlations between NPP and suitability for each species and each of the niche-based model techniques, with and without incongruences (previously identified in the robust regression). In general, correlation between NPP and suitability increased when incongruent VFs were removed (Table 3.3, values outside brackets). In the case of *Quercus ilex*, correlations increase and change from negative to positive when incongruences are not taken into account. On the other hand, *Pinus halepensis* showed high negative correlations between NPP and suitability, confirming the results observed in the robust regression.

	<i>Quercus ilex</i>	<i>Pinus sylvestris</i>	<i>Pinus halepensis</i>
ENS	0.30 (-0.15)	0.56 (0.26)	-- (-0.51)
ANN	0.33 (-0.11)	0.47 (0.19)	-- (-0.49)
CTA	0.27 (-0.17)	0.58 (0.30)	-- (-0.48)
GAM	0.31 (-0.29)	0.59 (0.29)	-- (-0.59)
GBM	0.31 (-0.17)	0.58 (0.35)	-- (-0.53)
GLM	0.35 (-0.23)	0.59 (0.30)	-- (-0.56)
MARS	0.29 (-0.12)	0.46 (0.19)	-- (-0.50)
RF	0.26 (-0.11)	0.51 (0.24)	-- (-0.41)

Table 3.3 Correlation between modeled NPP and suitability for different niche-based model techniques (see Table 1). Values in brackets correspond to all simulated forests, whereas values outside correspond to the congruence region (outliers from the robust regression removed). All correlations are significant ($p < 0.05$). *Pinushalepensis* present an overall incongruent pattern (negative robust relationship), thus no congruence region has been analyzed.

As expected, the most widespread species in the study region (*Quercus ilex*) showed less correlation (0.3 on average) than species with more restricted distribution (*Pinus sylvestris* and *Pinus halepensis*; 0.54 and -0.51 on average, respectively). In general, traditional approaches to regression techniques (GAM, GLM) tend to score higher correlations (Table 3.3) than classic machine-learning approaches (CTA, RF, ANN). Machine-learning techniques may tend to fit data

better, but they may easily assign a low score to absences in suitable conditions. In contrast, general regression models present smoother response functions, resisting the influence of absences in suitable conditions if they do not constitute the majority of absences. However, this may be highly depending on the way these models have been specified and also exceptions to this pattern are also present. For instance, GBM, although defined as a data mining technique, nearly performs at a level as high as that of traditional regressions. We recommend further testing on such behavior, using virtual species in order to control for other factors.

We also further analyzed the binary transformation applied by using a threshold measure based on TSS. As our data were not normally distributed, we performed analysis of variance using a non-parametric test (Kruskal-Wallis) for the variable NPP comparing two categories (suitable/unsuitable conditions). The results differ slightly from one model to another, but all of them support significant differentiation between suitable and unsuitable simulated forests in terms of NPP (see Appendix 2.3).

3.4 Discussion

In *Quercus ilex* and *Pinus sylvestris*, suitability estimates from climatic variables and current species distribution are positively related to ecophysiological performance derived from process-based modelling, supporting the hypothesis that the distribution of these species in this territory climate is broadly linked to physiology. Accordingly, *Quercus ilex* and *Pinus sylvestris* MLL and LAI have a positive relationship with climatic suitability for a broad area in the Iberian Peninsula. Nevertheless, both species present isolated regional discrepancies between the two modelling approaches. I^+ dominate over I^- in both species, demonstrating the availability of unoccupied potential environments for growth. Conversely, in *Pinus halepensis*, the relationship is the opposite: high NPP appears at sites with low suitability, meaning that factors other than climate would be more relevant than climate itself to the explanation of its distribution, at least on our scale of analysis.

Our results highlight the fact that the North of the Iberian Peninsula represents a potential region for the growth of forests of the three species studied. I^+ discrepancies are identified in that region for the two species showing general congruence (*Quercus ilex* and *Pinus sylvestris*), but the region also presents high growth values for the species showing no general congruence (*Pinus halepensis*). Regions of I^+ indicate physiological suitability, so other factors should be shaping current distributions. Biotic interactions have been demonstrated to strongly affect species distribution on different scales (Hampe, 2004; Araújo & Luoto, 2007). On our scale of analysis, competition from species (i.e., trees) using similar resources (particularly water and light) constitutes a fundamental ecological factor in a multispecies environment, especially in favourable abiotic conditions (Meier et al., 2010). In fact, Gomez-Aparicio et al., (2011) observed that competition is more important than climate for contemporary tree growth in the Iberian Peninsula, specifically for the same species studied here and with similar data.

Anthropogenic factors also influence vegetation distribution and they have long affected vegetation in the Iberian Peninsula (Perez-Obiol et al., 2011), and the geographical distributions of errors in NBM should be partly explained by human land use (Hanspach et al., 2011). In addition, differences in dispersal rates and past historical perturbations may also be causing such I^+ . For instance, the Last Glacial Maximum influenced many tree distributions and there is some evidence to show that it significantly explains tree richness in Europe (Svenning et al., 2010).

I^- was apparent only in the case of *Pinus sylvestris*. This incongruence was limited only to 0.5% of the VFs and was located in the South-East mountain ranges of the Iberian Peninsula, where a small population of the species still exists. We hypothesise four possible causes for this incongruence. First, the role of temporal variability in climatic drivers may have been underestimated. NBM use long-term averaged climatic factors to characterise the niche and suitable conditions, whereas our growth model simulated transient climate variables. Extreme climatic events, such as droughts in the Mediterranean (Lloret et al., 2007), may be important drivers for species distributions and *Pinus sylvestris* could be more vulnerable to such extreme events because its distribution in the Iberian Peninsula corresponds to its southern European limit (Galiano et al., 2010). Thus, NBM may gain from including such climatic episodes (Zimmermann et al., 2009). Second, the PBM used does not take species plasticity into account, whereas current species plasticity is considered, indirectly, by

presence records in NBM (McMahon et al., 2011). This could lead to a situation in which locations deemed highly suitable by the NBMs experience significantly less net productivity than that expected by the PBM. Third, biological interactions, particularly facilitation, may also play an important role blurring species growth from climate signal. Finally, species distribution is not static and may reflect a response to large-scale perturbations in the past (e.g. refuges during the Last Glacial Maximum) or management practices (e.g. plantations).

Many statistical techniques could be used to identify outliers between the outcomes of different models. Robust statistics have proved resistant to outliers and have been widely used in many fields (Maronna & Yohai, 2006), although other techniques, such as the robust generalised linear model (Azadeh & Salibian-Barrera, 2011), could potentially improve such analysis. Inevitably, the use of different techniques will affect results and therefore the determination of geographical incongruence areas. One way to minimise this is the use of an ensemble of these classification techniques to reduce uncertainty. Another interesting approach could be to avoid sharp classifiers in favour of soft or fuzzy ones (Zimmermann, 2000), allowing for a determination of a degree of congruence instead of two mutual exclusive categories.

Our result showed that the choice of NBM technique influences the magnitude of the growth-suitability relationship that is obtained (Table 3.3), although there is a common pattern in which widespread species (e.g. *Quercus ilex*) show lower correlation than regionally restricted species (e.g. *Pinus sylvestris* and *Pinus halepensis*). These results are in agreement with other studies showing that suitable/unsuitable habitats are more difficult to differentiate using NBM for widespread species rather than for restricted-climatic species (Grenouillet et al., 2011). Despite this variability between techniques, we were able to differentiate between suitable and unsuitable habitats. Nevertheless, we did not analyse different threshold criteria, which could potentially change these results (Nenzen & Araujo, 2011). All these sources of uncertainty suggest that the use of NBM ensembles obtains more robust results than the use of a single technique (Araújo & New 2007).

Most importantly, the species distribution range used for calibrating NBM affects the scale of suitability values, particularly if the complete range of the species is not used (Thuiller, 2004). In our

study, the inclusion of the whole species range would probably facilitate the determination of a general congruence between physiological performance and suitability. However, in the case of *P.sylvestris* it is reasonable to assume to apply only Iberian populations (Benito-Garzon 2006) which differ morphologically and genetically from the European ones (Ruby, 1967; Prus-Glowacki et al., 2003). In the remaining cases (*Q.ilex* ; *P.halepensis*), we argue that the Iberian peninsula encompasses the variability inherent to the Mediterranean climate to expect any significant change that would not support our conclusions. Unfortunately, the inclusion of the whole range of distribution implies the use of very different inventory methods in European countries, which may lead to biased results (Edwards et al., 2006), but most importantly such species distribution information is unavailable for some Mediterranean countries.

PBM results are contingent on model structure and the parameters applied (Raupach et al., 2005; Keenan et al., 2011b). Although PBMs have been widely tested, significant differences emerge between models when applied both at site level and on the regional scale (e.g., Schwalm et al., 2010). The PBM used in this study, GOTILWA+, has been rigorously tested at site level on various Mediterranean sites (Keenan et al., 2009a) but uncertainties associated with spatial scaling outside the domain of testing still remain. Nevertheless, general geographical gradients are often well captured by PBMs (Schwalm et al., 2010) and, in fact, macroscales often reveal a higher degree of congruence between climate, physiology and NBM (Helaouët & Beaugrand, 2009). The slope of the response between PBM productivity and NBM suitability is therefore likely to be subject to uncertainty due to potential PBM structural errors and parameter uncertainty. We would argue, however, that the general trend identified reflects the large-scale sensitivity of the PBM to climate and is thus relatively robust.

Although the PBM used is temporally dynamic, other potentially important dynamic processes are either static or not included. For example, spatially variable dynamics such as local adaptation and species plasticity could be significant when considering a species over its entire distribution. In the Mediterranean basin, phenotypic variation is known to allow forest tree species to adapt to macro-environmental gradients (Volis et al., 2002). The existing adaptation of populations to local conditions could bias results in the current study, for both PBMs and NBMs (Hampe, 2004). Unfortunately, we are far from a comprehensive understanding of possible species-specific

adaptation capacities. Recruitment dynamics also play a large role in species distributions. Seed dispersal, germination and development are subject to bioclimatic limits that are not included in the PBM used.

Generating a full characterization of PBM uncertainty is an onerous computational task, particularly when regional applications are taken into consideration (Zaehle et al., 2005). Without an associated estimate of error, the values presented from the PBM should only be considered as qualitative (Raupach et al., 2005). We argue however, that the general spatial gradients in PBM fields of NPP remain informative when compared with results from NBMs.

Far from raising another source of uncertainty and a new critique of NBM, the observed discrepancies throw new light on the hierarchy of factors such as human intervention or climate history contingency that constrain tree species distribution in a Mediterranean context. Accordingly, we have been able, for the first time, as far as we know, to map congruence and incongruence regions between NBM-generated habitat suitability and a PBM physiology-based performance. In the regions of discrepancy, other models dealing with processes linked to the species interactions or the geographical dynamics are strongly needed. We suggest that insights of species distribution may be discovered by the combination of such methods rather than claiming the superiority of one another (Dormann et al., 2012).

3.5 References

- Ainsworth EA, Leakey ADB, Ort DR, Long SP (2008) FACE-ing the facts: inconsistencies and interdependence among field, chamber and modelling studies of elevated [CO₂] impacts on crop yield and food supply. *New Phytologist*, **179**, 5-9.
- 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 MB, Pearson RG (2005) Equilibrium of species distributions with climate. *Ecography*, **28**, 693-695.
- Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, **16**, 743-753.
- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42-47.
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101-118.
- Austin MP, Smith TM, Niel KP, Wellington AB (2009) Physiological responses and statistical models of the environmental niche: a comparative study of two co-occurring Eucalyptus species. *Journal of Ecology*, **97**, 496-507.
- Austin MP, Van Niel KP (2011) Improving species distribution models for climate change studies: variable selection and scale. *Journal of Biogeography*, **38**, 1-8.
- Azadeh A, Salibian-Barrera M (2011). An outlier-robust fit for Generalized Additive Models with applications to disease outbreak detection. *Journal of the American Statistical Association*, **106**, 719-731.
- Ball JT, Woodrow IE, Berry JA (1987) A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In: *Progress in Photosynthesis Research* (ed. Biggins J), pp. 221–224. Martinus- Nijhoff Publishers, Dordrecht, the Netherlands.

- Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012), Selecting pseudo-absences for species distribution models: how, where and how many?. *Methods in Ecology and Evolution*. In press.
- Benito Garzón M, Blazek R, Neteler M, Sanchez de Dios R, Sainz Ollero H, Fulanello C (2006). Predicting habitat suitability with machine learning models: The potential area of *Pinussylvestris* L. in the Iberian Peninsula. *Ecological Modelling*, **197**, 383-393.
- Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and Regression Trees. Chapman & Hall, New York.
- Breiman L (2001) Random forests. *Machine Learning*, **45**, 5–32.
- Buckley L (2008). Linking traits to energetics and population dynamics to predict lizard ranges in changing environments. *American Naturalist*, **171**, E1–E19.
- Bugmann H (2001) A Review of Forest Gap Models. *Climatic Change*, **3**, 259.
- Chuine I, Beaubien EG (2001) Phenology is a major determinant of tree species range. *Ecology Letters*, **4**, 500-510.
- Campbell GS (1986) Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution. *Agricultural and Forest Meteorology*, **36**, 317-321.
- Dai YJ, Dickenson RE, Wang YP (2004) A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *Journal of Climate*, **17**, 2281–2299.
- Dormann CF (2007) Promising the future? Global change projections of species distributions. *Basic and Applied Ecology*, **8**, 387-397.
- Dormann CF, Schymanski S J, Cabral J, et al. (2012), Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*. In press
- Edwards J, Thomas C, Cutler DR, Zimmermann NE, Geiser L, Moisen GG (2006) Effects of sample survey design on the accuracy of classification tree models in species distribution models. *Ecological Modelling*, **199**, 132-141.
- Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677-697.

- Falk W, Mellert KH (2011) Species distribution models as a tool for forest management planning under climate change: risk evaluation of *Abies alba* in Bavaria. *Journal of Vegetation Science*, **22**, 621-634.
- Farquhar GD, Caemmerer S, Berry J. (1980) A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta*, **149**, 78-90.
- Fielding AH (2002). What are the appropriate characteristics of an accuracy measure? In: Predicting Species Occurrences: Issues of Accuracy and Scale (eds Scott, J.M., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A. & Samson, F.B.). Island Press, Covelo, CA, pp. 271–280.
- Franklin J (2010) Mapping species distributions :spatial inference and prediction, Cambridge University Press, Cambridge.
- Friedman J (1991) Multivariate adaptive regression splines. *Annals of Statistics*, **19**, 1–1141.
- Friedman JH (2001) Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, **29**, 1189-1232.
- Galiano L, Martinez-Vilalta J, Lloret F (2010) Drought-induced multifactor decline of Scots pine in the Pyrenees and potential vegetation change by the expansion of co-occurring oak species. *Ecosystems*, **13** ,978-991.
- Gomez-Aparicio L, Garcia-Valdes R, Ruiz-Benito P, Zavala MA (2011) Disentangling the relative importance of climate, size and competition on tree growth in Iberian forests: implications for forest management under global change. *Global Change Biology*, **7**, 2400-2414.
- Gracia CA, Tello E, Sabate S, Bellot J (1999) GOTILWA: an integrated model of water dynamics and forest growth. In: Ecology of Mediterranean Evergreen Oak Forests (eds Roda F, Retana J, Gracia CA, Bellot J), pp. 163–179. Springer-Verlag, Berlin.
- Graham CH, Ron SR, Santos JC, Schneider CJ, Moritz, C. (2004) Integrating phylogenetics and environmental niche models to explore speciation mechanism in *Dendrobates* frogs. *Evolution*, **58**, 1781-1793.
- Grenouillet G, Buisson L, Casajus N, Lek S (2011) Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography*, **34**, 9-17.

- Hampe A (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography*, **13**, 469-471.
- Hanspach J, Kühn I, Schweiger O, Pompe S, Klotz S (2011) Geographical patterns in prediction errors of species distribution models. *Global Ecology and Biogeography*, **20**, 779-788.
- Hartigan JA, Wong MA (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **28**, 100–108.
- Hastie TJ, Tibshirani R (1990) Generalized Additive Models. Chapman & Hall, New York.
- Helaouët P, Beaugrand G (2009) Physiology, Ecological Niches and Species Distribution. *Ecosystems*, **12**, 1235.
- Jeltsch F, Moloney KA, Schurr FM, Kikvidze Z, Schwager M (2008) The state of plant population modelling in light of environmental change. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 171-189.
- Kearney M (2006) Habitat, environment and niche: what are we modelling? *Oikos*, **115**, 186-191.
- Kearney M, Phillips BL, Tracy CR, Christian, KA, Betts G, Porter WP (2008) Modelling species distributions without using species distributions: the cane toad in Australia under current and future climates. *Ecography*, **31**, 423-434.
- Keenan T, Garcia R, Friend AD, Zaehle S, Gracia C, Sabate S (2009a) Improved understanding of drought controls on seasonal variation in Mediterranean forest canopy CO₂ and water fluxes through combined in situ measurements and ecosystem modelling. *Biogeosciences*, **6**, 1423–1444.
- Keenan T, Niinemets U, Sabate S, Gracia C, Peñuelas J. (2009b) Process based inventory of isoprenoid emissions: current knowledge, future prospects and uncertainties. *Atmospheric Chemistry and Physics*, **9**, 4053–4076.
- Keenan T, Niinemets U, Sabate S, Gracia C, Peñuelas J (2009c) Seasonality of monoterpene emission potentials in *Quercus ilex* and *Pinus pinea*: implications for regional VOC emissions modelling. *Journal of Geophysical Research*, **114**, D22202.

- Keenan T, Serra-Diaz JM, Lloret F, Ninyerola M, Sabate S (2011a) Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO2 matters! *Global Change Biology*, **17**, 565-579.
- Keenan TF, Carbone M, Reichstein M, Richardson AD (2011b) The model-data fusion pitfall: assuming certainty in an uncertain world. *Oecologia*, **167**, 587-597.
- Leuning R, Kelliher FM, De Pury DGG, Schulze ED (1995) Leaf nitrogen, photosynthesis, conductance and transpiration: scaling from leaves to canopies. *Plant, Cell and Environment*, **18**, 1183–1200.
- Lischke H, Zimmermann NE, Bolliger J, Rickebusch S, Löffler TJ (2006) TreeMig: A forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. *Ecological Modelling*, **199**, 409-420.
- Lloret F, Lobo A, Estevan H, Maisongrande P, Vayreda J, Terradas J (2007) Woody plant richness and NDVI response to drought events in Catalanian (northeastern Spain) forests. *Ecology*, **88**, 2270-2279.
- Larcher W (2000) Temperature stress and survival ability of Mediterranean sclerophyllous plants. *Plant Biosystems*, **134**, 279-295.
- Maronna RA, Douglas Martin R, Yohai VJ (2006) Robust statistics: theory and methods. John Wiley & Sons (Chichester).
- McCullagh P, Nelder J.A. (1989) Generalized Linear Models. Chapman & Hall, New York.
- McMahon SM, Harrison SP, Armbruster WS, et al. (2011) Improving assessment and modelling of climate change impacts on global terrestrial biodiversity. *Trends in Ecology & Evolution*, **26**, 249-259.
- Meier ES, Kienast F, Pearman PB, et al. (2010) Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, **33**, 1038-1048.
- Ministerio de Agricultura PyA (2007) Tercer Inventario Forestal Nacional. Ministerio de Agricultura PyA, Madrid, Spain.
- Morales P, Sykes MT, Prentice IC et al. (2005) Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes. *Global Change Biology*, **11**, 1–23.

- 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.
- Nenzén H, Araújo M (2011) Choice of threshold alters projections of species range shifts under climate change. *Ecological Modelling*, **222**, 3346-3354.
- Ninyerola M, Pons X, Roure JM (2000) A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology*, **20**, 1823-1841.
- Ninyerola M, Pons X, Roure JM (2007a) Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. *International Journal of Climatology*, **27**, 1231-1242.
- Ninyerola M, Pons X, Roure JM (2007b) Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. *Theoretical and Applied Climatology*, **89**, 195-209.
- Ninyerola M, Serra-Díaz JM, Lloret F. 2010. Topo-climatic Suitability Atlas of Woody Plants. Map server. Universitat Autònoma de Barcelona. URL:<http://www.opengis.uab.cat/IdoneitatPI/index.html>
- Pelkonen P, Hari P (1980) The dependence of the springtime recovery of CO₂ uptake in Scots pine on temperature and internal factors. *Flora*, **169**, 398–404.
- Perez-Obiol R, Jalut G, Julia R, Pelachs A, Iriarte MJ, Otto T, Hernandez-Beloqui B (2011) Mid-Holocene vegetation and climatic history of the Iberian Peninsula. *The Holocene*, **21**, 75.
- Prus-Glowacki W, Stephan BR, Brujas E, Alia R, Marciniak A (2003) Genetic differentiation of autochthonous populations of *Pinus sylvestris* (Pinaceae) from the Iberian Peninsula. *Plant Systematics and Evolution*, **239**, 55–66.
- R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.
- Randin CF, Engler R, Normand S, et al. (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, **15**, 1557-1569.

- Raupach MR, Rayner PJ, Barrett DJ, et al. (2005) Model data synthesis in terrestrial carbon observation: methods, data requirements and data uncertainty specifications. *Global Change Biology*, **11**, 378-397.
- Renaud O, Victoria-Feser M (2010) A robust coefficient of determination for regression. *Journal of Statistical Planning and Inference*, **140**, 1852-1862.
- Rickebusch S, Thuiller W, Hickler T, et al. (2008) Incorporating the effects of changes in vegetation functioning and CO₂ on water availability in plant habitat models. *Biology Letters*, **4**, 556-559.
- Ruby JL (1967) The correspondence between genetic, morphological and climatic variation patterns in Scotch Pine. *Silvae Genetica*, **16**, 50–56.
- Ripley BD (1996) Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge, UK.
- Rousseeuw P, Croux C, Todorov V, et al. (2011). robustbase: Basic Robust Statistics. R package version 0.7-3. URL <http://CRAN.R-project.org/package=robustbase>
- Scheller RM, Mladenoff DJ (2007) An ecological classification of forest landscape simulation models: tools and strategies for understanding broad-scale forested ecosystems. *Landscape Ecology*, **22**, 491–505.
- Sitch S, Smith B, Prentice IC, et al. (2003) Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, **9**, 161-185.
- Svenning J, Skov F (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, **7**, 565-573.
- Svenning J, Fitzpatrick MC, Normand S, et al. (2010) Geography, topography, and history affect realized-to-potential tree species richness patterns in Europe. *Ecography*, **33**, 1070-1080.
- Thomas CD, Cameron A, Green RE, et al. (2004) Extinction risk from climate change. *Nature*, **427**, 145-148.
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD : a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369-- 373.

- Thuiller W (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020-2027.
- Thuiller W, Lavorel S, Araújo MB, Sykes MT, Prentice IC (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 8245-8250.
- Thuiller W, Richardson DM, Pyek P, Midgley GF, Hughes GO, 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.
- Volis S, Mendlinger S, Turuspekov Y, Esnazarov U (2002) Phenotypic and allozyme variation in Mediterranean and desert populations of wild barley, *Hordeum spontaneum*. *Evolution*, **56**, 1403–1415.
- Wang YP, Leuning R (1998) A two-leaf model for canopy conductance, photosynthesis and partitioning of available energy I: model description and comparison with a multi-layered model. *Agricultural and Forest Meteorology*, **91**, 89–111.
- Zaehle S, Sitch S, Smith B, Hatterman F (2005) Effects of parameter uncertainties on the modelling of terrestrial biosphere dynamics. *Global Biogeochemical Cycles*, **19**, GB3020.
- Zimmermann HJ (2000). *Practical Applications of Fuzzy Technologies*. Springer.
- Zimmermann NE, Yoccoz NG, Edwards TC, et al. (2009) Climatic extremes improve predictions of spatial patterns of tree species. *Proceedings of the National Academy of Sciences*, **106**, 19723-19728.

4. ON THE IMPORTANCE OF (PHYSIOLOGICAL) MECHANISMS IN CLIMATE CHANGE PROJECTIONS

STUDY: 'Predicting the future of Mediterranean forests with niche- and process-based models: CO₂ matters!'²

4.1 Introduction

Detailed knowledge of species' ecological and geographic distributions is fundamental for conservation planning and forecasting (Ferrier, 2002; Funk & Richardson, 2002; Rushton et al., 2004), for understanding ecological and evolutionary determinants of the spatial patterns of biodiversity (Ricklefs, 2004), and the potential response of these distributions to future climatic change (e.g. Thomas et al., 2004; Araújo et al., 2005a, b; Thuiller et al., 2005). This is of particular importance in the Mediterranean region, which has a high diversity of environments and harbors Europe's greatest diversity of vegetation and fauna (Cowling et al., 1996). This region is not only a biodiversity hotspot (Underwood et al., 2009), it has also been identified as a climatic change hotspot (Giorgi, 2006) because (1) climate projections consistently project significant increases in temperature, and decreases in precipitation in the Mediterranean basin (Gibelin & Deque, 2003; Giorgi et al., 2004) and (2) such potential change may have a large effect on current Mediterranean forests and the related ecosystem service supply (Schröter et al., 2005).

Models are the most feasible and efficient way to assess spatial biodiversity and responses to climatic drivers over large spatial and temporal scales (Thuiller, 2007). Species-specific models fall broadly into two categories: empirical niche-based or habitat models and process-based models

² This study was led by Dr. Keenan, who performed physiological simulations. JM Serra-Diaz performed the analysis and described results for correlative ecological niche models. Conclusions are drawn from the interaction of the two researchers with great input from other co-authors. Manuscript available at: *Global Change Biology*, 17, 565-579.

(see Kearney, 2006). These contrasting methodologies, however, often give conflicting results (Thuiller et al., 2008).

Also known as ecological species distribution models, bioclimatic envelopes, or habitat models, niche-based models are by far the most commonly used method for predicting species geo-climatic distributions. Such models typically use a variety of correlative methods to provide detailed predictions of distributions by relating presence or abundance of species to gradients of observed environmental predictors. As such, niche-based models are used extensively and have provided researchers with an innovative tool to explore diverse questions in ecology and conservation (see Peterson, 2007). In particular, it has become common to use such models to assess potential distribution responses to future climate scenarios (e.g., Bakkenes et al., 2002; Araújo et al., 2004; Skov & Svenning, 2004; Thomas et al., 2004; Thuiller et al., 2005; Gomez-Mendoza & Arriaga, 2007; Thuiller, 2010), using sophisticated interpolation of climate data (e.g., Hijmans et al., 2005). One of the main advantages of niche-based models is their relative simplicity, making it straightforward to develop species-specific models, which make use of the large data sets available (e.g., Forest inventories, regionalized climate).

For terrestrial vegetation, the term 'process-based model' incorporates a broad range of methodologies for describing eco-physiological processes, from purely empirical relationships to mechanistic descriptions based on physical laws. Various types of process-based models are used and under development, such as gap models (Pacala et al., 1993; Bugmann, 2001), landscape models (Lischke et al., 2006; Scheller & Mladenoff, 2007), fitness-based models (Chaine & Beaubien, 2001), or sophisticated 'hybrid' dynamic vegetation models (e.g., Sitch et al., 2008), which focus on achieving a balance between realism, accuracy and complexity. The suite of available models represents a range from very detailed species-specific models which describe stand structure and hourly plant physiological processes (i.e. coupled photosynthesis, respiration, and water balance), to general models based on fitness probability matrices. A process-based model can potentially allow for the highlighting of processes involved in range shifts or extinction. To date, various process-based approaches have been developed for predicting species distributions (see Jeltsch et al., 2008 for a review). These have been explicitly empirical, and link simple indexes of survival and reproduction with impacts of frost, drought, and windthrow to

produce a presence–absence indicator. The use of these empirical models to make predictions of species range shifts is rare for species ranges at the regional scale (Hijmans & Graham, 2006; Jeltsch et al., 2008).

Many mechanistic process-based model studies, supported by experimental campaigns such as the FACE project (Ainsworth & Long, 2005; Ainsworth & Rogers, 2007), as well as growth and yield surveys, suggest that global warming will have a positive impact on forest productivity (van der Meer et al., 2002; Nigh et al., 2004; Norby & Luo, 2004; Briceño-Elizondo et al., 2006; Gaucharel et al., 2008), due to the direct fertilization effect of increased CO₂ and indirect effects such as lengthening of the growing period [but see contrary examples such as Zierl & Bugmann (2007)]. Results vary among experimental systems, especially when considering potential acclimation (Körner, 2006) and nutrient limitation (Zaehle et al., 2010), and remain the focus of much study. On the other hand, results from statistical niche-based models are supported by a growing body of ecological literature that suggests that the narrow climatic adaptation of many tree species may lead to many populations being poorly suited to their environment, resulting in large alteration to potential distributions towards the end of the 21st Century (Davis & Shaw, 2001; Iverson & Prasad, 2001; Thomas et al., 2004; St Clair & Howe, 2007; Benito-Garzon et al., 2008). Within studies, different modeling methods yield highly divergent predictions, even when spatial assessments of model accuracy appear excellent (Araújo et al., 2005a, b; Kharouba et al., 2009).

It is of great importance to develop several methods independently and to compare (for the same species and under the same scenarios) their predictions in order to identify both robust results and model inadequacies (Beaumont et al., 2007). Such cross comparisons may provide information on which policy makers and stakeholders can rely. Yet, despite the uncertainty generated by contrasting experimental results, the variety of modeling approaches available (Jeltsch et al., 2008), and studies that have highlighted differences between niche-based modeling approaches (Elith et al., 2006; Hijmans & Graham, 2006), niche-based model predictions are rarely compared against other modeling approaches. Of particular relevance to this study, the models used in previous comparisons (e.g., Hijmans & Graham, 2006; Jeltsch et al., 2008; Coops et al., 2009; Morin & Thuiller, 2009) have not described the indirect effect of CO₂ driven forest productivity on the suitability of a site (but see Rickebusch et al., 2008).

This paper has three purposes. First, we consider the effectiveness of an empirically derived multi-niche-based model ensemble, applying the *biomod-r* package (Thuiller et al., 2009) with regionalized present-day (1950–1998) climate, to predict the distributions of three forest species in continental Spain. These species are presently distributed along a gradient from drier (*Pinushalepensis*), to mesic (*Quercus ilex*) and moister conditions (*P. sylvestris*). Second, we assess potential future climate driven changes in current forest stands using both the niche-based approach and a mechanistic process-based forest growth model (GOTILWA+). Third, we use the process-based simulations of projected future forest productivity (with and without CO₂ fertilization) to identify possible processes responsible for the large differences observed in future projections from the two modeling approaches. We then suggest possible means by which to improve future model efforts.

4.2 Material and methods

Selected region and study species

We focused on the region of continental Spain, which contains a large altitudinal gradient (sea level to 3500 m) and a mosaic of different climates (from semiarid climates to Mediterranean and humid Atlantic climates). The Third Spanish National Forest Inventory (Ministerio de AgriculturaPyA, 2007) surveyed the forested surface of the Spanish Iberian Peninsula (492 173 km²) with an approximate density of 1 plot km². Each of the resulting 89 365 circular plots was located in the field by giving its Universal Transverse Mercator (UTM 30N) coordinates. We extracted the presence/absence data for three species with distinct topoclimatic distributions: two tree species (*Q. ilex*, *P. halepensis*) commonly found in Mediterranean forests of the study region, and one species (*P. sylvestris*) which is found at the most southern limits of its distribution (Figure 4.1).

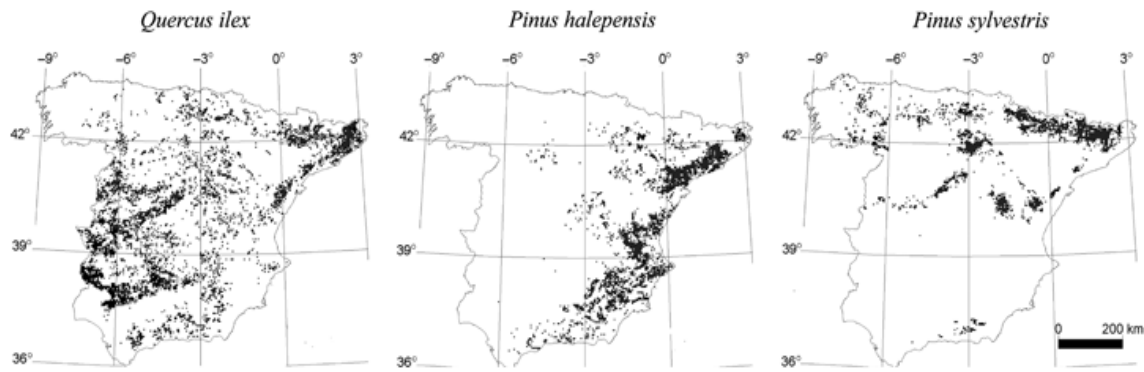


Figure 4.1 Current distribution of *Q.ilex*, *P.halepensis* and *P.sylvestris* in continental Spain.

Climatic data

From 1900 to 2000, a reconstructed climatic data time series based on the CRU05 (1901–2000) monthly data set (New et al., 1999) was used, with global atmospheric concentrations of CO₂ from 1901 to 2000 obtained from the Carbon Cycle Model Linkage Project (McGuire et al., 2001).

We applied the climatic projection for period 2001–2100 generated by the HadCM3 global circulation model using the A1 scenario (IPCC WGI, 2007) as an indicator for the effect of possible future climate change on species distributions and productivity. The HadCM3 model with the A1 scenario uses an estimated increase in atmospheric CO₂ to 810 ppm by 2080, with an associated increase in temperature of 3.1 °C by 2080 for the area included in this study (in comparison with the average temperatures for the 1960–1990 period), and a slight decrease in precipitation.

The present-day regionalization was created by GIS modeling from ground data (1950–1998; 1068 thermometric and 1999 pluviometric meteorological stations) from the Spanish weather monitoring system (National Weather Institute; <http://www.aemet.es>). The regionalization method chosen was multiple regression with residual correction (spatially interpolated using inverse distance weighting in the case of mean temperatures or splines in the case of precipitation). The climate predictors were altitude, latitude, continentality (linear or quadratic distances to Mediterranean, Atlantic and Cantabric coasts) and potential global solar radiation; all of them derived from a DEM (Digital Elevation Model). A holdout crossvalidation, using a fitting set (60%) and a testing set (40%), was applied to compute the RMSE for the monthly data (between 0.8 and 1.6 °C in the case of temperature and between 6 and 20 mm in the case of precipitation). See Ninyerola et al. (2007a, b) for more details on the methodology used. The resulting maps have a

200 m spatial resolution, although in this study we have generalized the matrix into a 1 km grid for computational purposes.

Future regionalized climate was obtained using an approximation based on differences between the past climate (CRU) and the climate projection from the HadCM3 model using the A1 storyline, thus combining the predictive information of the GCM with the topoclimatic data provided by ground stations.

Niche-based models

We performed the projections using nine different and widely used niche-based modeling techniques, within the BIOMOD computational framework (Thuiller, 2003; Thuiller et al., 2009), as outlined in Table 4.1.

All models used in this study need information about presences and absences to be able to determine suitable conditions for a given species. Pseudo-absences were randomly selected from stands at least 10 km distant to a presence where the target species was not recorded in the Forest Inventory. The number of pseudo-absences and presences were equaled in order to keep prevalence constant. A holdout crossvalidation has been used to evaluate the models: 80% of the presence-absence stands have been labeled as the training set and the remaining 20% as the testing set. The number of model evaluation runs was set to three with a repetition of two pseudo-absence combinations on each run, resulting in a total number of six evaluation runs per model.

The models outlined predict suitability values between 0 and 1 at each site. We applied a binary transformation (absence or presence, 0 or 1) by assigning a threshold from which we can consider a species present or absent, using the True Skill Statistic (TSS). This threshold represents the optimum correct classification of both presences and absences within the evaluation data. However, the accuracy of each model was calculated using both the AUC [area under the curve, from receiver operating characteristics curve (Swets, 1988)] and TSS (Allouche et al., 2006).

To constrain model uncertainty, averaging of model predictions (giving the same weight to all predictions) can be implemented to derive a consensus prediction; an alternative is to combine models using some form of weighting (e.g. using PCA score value, Thuiller et al., 2003; Araújo &

Guisan, 2006). There is a wide range of approaches to do this (see Araújo & New, 2007 for a review). In this study, each model is given a weight in the ensemble of forecasts depending on their predictive accuracy using TSS. A decay factor of 1.6 in weights is set; that is, models are ordered in terms of TSS and the weight of each model in the forecast will be 1.6 times larger than the following. This procedure (i.e. a form of ‘stacking’) ensures accuracy-based discrimination between models. Further information on this procedure may be found in Thuiller et al. (2009).

Climatic and topographic uncorrelated variables (from more than 100 raw variables) are used to apply these niche-based models. For each target species, a different set of topoclimatic variables is chosen by evaluating the correlation matrices. Variables were chosen from those commonly shown to influence tree distribution.

The variables chosen for each species, and their mean relative importance over all niche-based models were:

Q. ilex – Summer minimum water availability (11%), mean winter water availability (31%), minimum of the mean winter temperature (23%), minimum of the mean summer temperature (12%), cost–distance to the sea (15%), slope (7%).

P. halepensis – Mean spring water availability (40%), minimum of the mean winter temperature (23%), cost–distance to the sea (23%), aspect (14%).

P. sylvestris – Summer minimum water availability (39%), mean winter water availability (8%), minimum of the mean winter temperature (35%), cost–distance to the sea (6%), slope (11%).

The relative importance of each variable is calculated based on the correlation between standard prediction and the prediction with a randomized variable, therefore estimating the influence of the randomized variable in the modeling (see Thuiller et al., 2009 for further details).

Model number	Abbreviation	Description	Reference
1	RF (random forest)	A machine-learning method – a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest.	Breiman (2001)
2	CTA (classification tree analysis)	A classification method – a 50-fold cross-validation to select the best trade-off between the number of leaves of the tree and the explained deviance.	Breiman et al. (1984)
3	GBM (generalized boosting model)	A machine-learning method – combines a boosting algorithm and a regression tree algorithm to construct an ‘ensemble’ of trees.	Ridgeway (1999)
4	MARS (multivariate adaptive regression splines)	A nonparametric regression method, mixing CTA and GAM.	Friedman (1991)
5	GAM (generalized additive model)	A regression method, with 4 degrees of freedom and a stepwise procedure to select the most parsimonious model.	Hastie & Tibshirani (1990)
6	MDA (mixture discriminant analysis)	A classification method – based on mixture models	Hastie & Tibshirani (1990)
7	GLM (generalized linear model)	A regression method, with polynomial terms for which a stepwise procedure is used to select the most significant variables.	McCullagh & Nelder (1989)
8	ANN (artificial neural networks)	A machine-learning method, with the mean of three runs used to provide predictions and projections.	Ripley (1996)
9	SER (surface range envelope)	A simple rectilinear envelope, that takes into account the whole range of conditions in which the species is present.	Busby (1991)

Table 4.1 Niche-based models used in this study

Ecosystem model GOTILWA+

GOTILWA+ (Growth Of Trees Is Limited by WAter), (Keenan et al., 2008, 2009a, b, c, 2010; <http://www.creaf.uab.es/GOTILWA+>) is a process-based terrestrial biogeochemical model of forest growth that has been developed in the Mediterranean region to explore how forests are influenced by water stress, tree stand structure, management techniques, soil properties, and climate (including CO₂) change.

GOTILWA+ does not predict the distribution of a species, but simulates tree growth, and the associated carbon and water fluxes for different tree species in different environments, thus reflecting a site-species specific ecophysiological suitability. The model treats monospecific stands, which can be even or uneven-aged. Individual trees are aggregated into 50 dbh (diameter at breast height) classes and calculations are performed for each class. Hourly ecosystem carbon and water fluxes are estimated using meteorological forcing. No bioclimatic limits are set in GOTILWA+, and indeed indirect bioclimatic limits can only be considered through the direct effect of climate on the physiological processes of the forest.

The GOTILWA+ model includes a two-leaf canopy photosynthetic submodel (Wang & Leuning, 1998; Dai et al., 2004). The photosynthesis submodel treats the C3 photosynthetic pathway. The canopy is divided into sunlit and shaded leaves, with the amount of intercepted diffuse and direct radiation depending on the time of the day, season, and the area of leaf exposed to the sun (Campbell, 1986). Foliage net assimilation rates are calculated using the Farquhar et al. (1980) photosynthesis model, with dependencies on intercepted quantum flux density, species-specific photosynthetic capacities, leaf temperature, and leaf intercellular CO₂ concentration. Stomatal conductance is calculated using the Leuning et al. model (Leuning et al., 1995) that is the advancement of the Ball et al. (1987) model. Net photosynthesis is scaled from the leaf to the canopy through the canopy microclimate model, to give canopy bulk gross primary production (GPP). Net primary production (NPP) is calculated as the balance of GPP less autotrophic respiration components and is defined as $NPP = A - R_f - R_w - R_r$, where A is the net assimilation rate per unit ground (=GPP–daytime leaf respiration), R_f is night respiration rate per ground unit area, R_w is respiration of nonleaf aerial plant tissues, R_r is respiration of root tissues. Model parameters were set to

species-specific values (as in Gracia et al., 1999; Kramer et al., 2002; Morales et al., 2005; Keenan et al., 2009a). Each tree cohort is represented with three carbon compartments: leaf, sapwood, and fine roots. Available mobile carbon is allocated to each, and maintenance respiration of each compartment is calculated as a function of temperature.

Water stress affects the photosynthesis–conductance coupling by directly reducing the photosynthetic potential through a nonlinear relation to soil water content (Keenan et al., 2009a). Phenology is temperature-dependent and accounted for in an updated version of the Pelkonen&Hari (1980) approach for calculating the seasonal variations in photosynthetic potential. GOTILWA+ has been validated and widely applied both in the Mediterranean region and the rest of Europe (see Kramer et al., 2002; Morales et al., 2005; Keenan et al., 2009a for validation exercises and Keenan et al., 2009b, c, 2010 for example applications).

Experimental setup

The niche-based model ensemble was used to calculate the per model suitability for each recorded stand (1 km² scale) for the two periods 1950–1998 and 2050–2080. A weighted mean model ensemble suitability was then calculated for each period. For GOTILWA+, simulations were run for each dominant occurrence of the three studied species for the period 1930–2100. Two experiments were considered for the GOTILWA+ model: (1) with increasing CO₂ concentrations as prescribed by the A1 climate scenario, (2) with CO₂ concentrations fixed constant at post 2000 levels from 2000 to 2100. We used modeled values of NPP as a pseudo-proxy for suitability, given that it reflects direct changes in temperature and soil water availability, and as well as more complex indirect effects of changes in phenological events, labile carbon pools, stand biomass and the associated maintenance, growth and turnover. Long- and short-term changes in NPP therefore can be used as a simple representation of the ‘health’ of a forest stand and may be correlated to changes in suitability values for a given species.

4.3 Results

We first assessed consistency in niche-based model predictions by measuring agreement between modeled present-day distributions and known presence and pseudo-absence of species (Table 4.2). The results showed a good predictive ability for observed distributions, with most mean AUC and TSS values within ranges of good predictive performance (Allouche et al., 2006). The Random Forest (RF) model performed consistently better across species, followed by the classification tree analysis (CTA) and generalized boosting model (GBM) models. Variability in performance between modeling techniques was high (Table 4.2), with mean TSS values varying by up to two times between models. The TSS statistic proved to be a more sensitive estimator of model predictive accuracy than the AUC statistic. In the case of TSS, each species weighted model ensemble proved to have higher predictive power than simply taking the average of all models, or even using the best model.

Two distinct groups were observed in the niche-based models: the first consisting of the three methods GAM, GLM, and MDA, and the second group comprising of the three methods MARS, GBM, and CTA. Three methods [artificial neural networks (ANN), RF, and surface range envelope (SRE)] with distinct predictions were observed (Figure 4.2). Although the majority of methods show an overall good performance of mean TSS across species, they vary in terms of TSS variance across species. The RF model proved to be the most stable across species due to its flexibility in contrast with SRE, a restricted envelope only based on maximum and minimum of the explanatory variables. ANN accounts for the greatest variance due to its performance for *Q. ilex* (Table 4.2).

MODELING TECHNIQUE	Quercus ilex			Pinushalepensis			Pinussylvestris		
	w	AUC	TSS	w	AUC	TSS	w	AUC	TSS
RF	0.381	0.974 (0.004)	0.845 (0.011)	0.381	0.962 (0.006)	0.805 (0.018)	0.381	0.989 (0.002)	0.911 (0.012)
CTA	0.238	0.931 (0.01)	0.768 (0.02)	0.238	0.924 (0.011)	0.750 (0.019)	0.22	0.963 (0.008)	0.876 (0.019)
GBM	0.149	0.947 (0.006)	0.770 (0.014)	0.149	0.954 (0.006)	0.779 (0.013)	0.16	0.984 (0.004)	0.881 (0.014)
MARS	0.093	0.929 (0.01)	0.714 (0.02)	0.073	0.943 (0.005)	0.750 (0.010)	0.1	0.981 (0.005)	0.881 (0.012)
GAM	0.031	0.893 (0.008)	0.643 (0.012)	0.078	0.942 (0.006)	0.755 (0.010)	0.055	0.978 (0.004)	0.873 (0.008)
MDA	0.054	0.890 (0.012)	0.668 (0.028)	0.021	0.01 (0.018)	0.699 (0.024)	0.02	0.871 (0.007)	0.861 (0.014)
GLM	0.032	0.890 (0.012)	0.645 (0.015)	0.036	0.933 (0.007)	0.737 (0.017)	0.035	0.976 (0.004)	0.867 (0.006)
ANN	0.01	0.754 (0.037)	0.475 (0.048)	0.016	0.903 (0.008)	0.706 (0.012)	0.021	0.963 (0.016)	0.847 (0.021)
SRE	0.013		0.528 (0.015)	0.009		0.436 (0.030)	0.009		0.683 (0.022)
Ensemble		0.958 (0.003)	0.961 (0.003)		0.969 (0.001)	0.976 (0.001)		0.990 (0.486)	0.990 (0.513)

Table 4.2 Assessment of the agreement between modeled and observed distributions for each niche-based model and species, and the resulting weights (w) used in the multimodel ensemble.

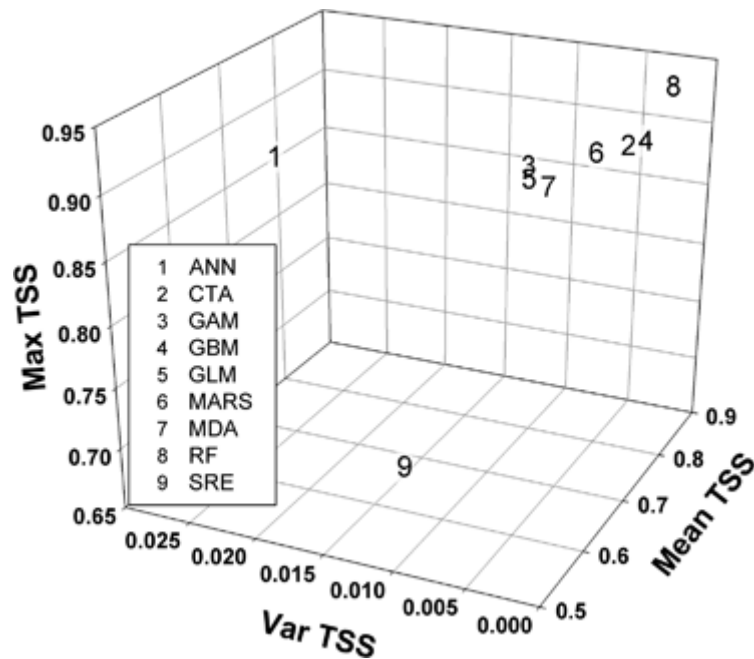


Figure 4.2 Variation of performance of each niche modeling technique

When predicting responses to climate change, a broad topoclimatic range of responses were observed (Figure 4.3), although all species showed the same general tendency. Model ensemble predictions of suitability showed large declines in suitability for each of the three species between the periods 2050–2080 and 1950–1980. *Q. ilex* stands were the largest affected by the applied climate change scenario (Figure 4.4), with 40.4% of current stand locations becoming unsuitable by the period 2050–2080. Although *Q. ilex* is a typical Mediterranean species, and relatively drought tolerant, its large topoclimatic distribution means that it is currently located in some areas which are predicted to be subject to high levels of climate change in the future. Thus, areas of its southern most range were the highest affected. Climate change induced decline of *P. halepensis* was not so severe, with the majority of sites (87.3%) maintaining a level of suitability that would permit the presence of the species. The multimodel ensemble also predicted an important decline in the presence of *P. sylvestris* (24%), though the species maintained a strong presence in most mountainous regions (e.g. the Pyrenees Mountains), resulting in much larger geographical variability than that observed for the other two species.

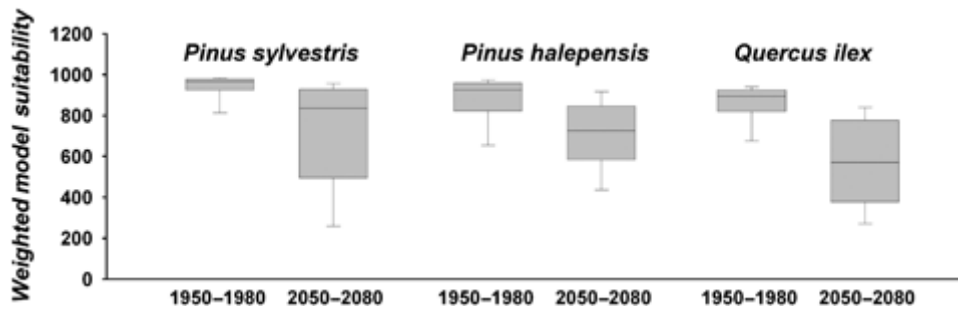


Figure 4.3 Changes in suitability in current forests

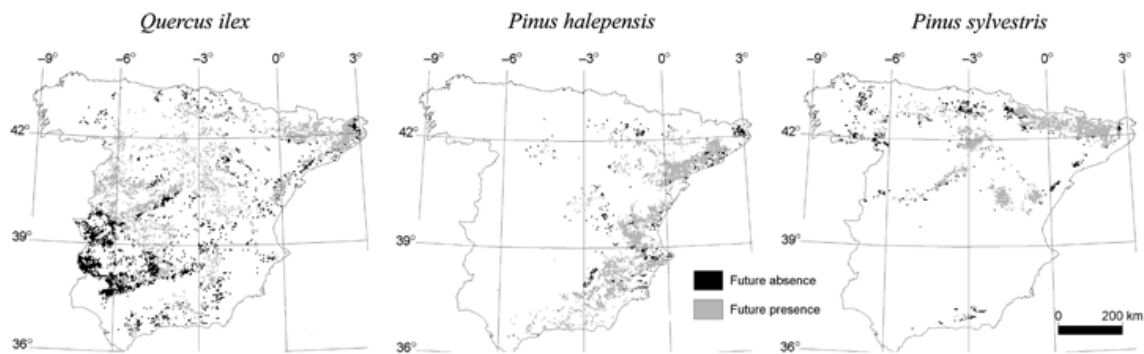


Figure 4. 4 Predicted future of current forests of *Quercus ilex*, *Pinus halepensis*, and *P. sylvestris* in continental Spain, as predicted by the multimodel ensemble for the period 2050–2080. Future absence relates to current forest stands in locations which are projected to be geoclimatically unsuitable by the period 2050–2080.

Results for estimates of change in current forests under future climates also demonstrate that the modeling technique used to define climate envelopes can have a very large impact on predictions (Table 4.3). Predictions for each of the three species (excluding SRE which presents a very low weight) varied in magnitude of predicted change. For example, for *Q. ilex* predicted losses of current habitat ranged from 5.6% to 46.9%. The other two species showed a lower range of model dependent variability (between 3% and 35%). Between-model variability across species was also very high, with models predicting between 17% (MARS model) and 28.3% (CTA model) loss of current habitat on average over the three species by the period 2050–2080.

	ANN	CTA	GAM	GBM	GLM	MARS	MDA	RF	SRE	Mean	SD
<i>Quercus ilex</i>	5.65	46.97	21.34	40.06	16.92	29.1	9.66	30.56	88.42	32.08	25
<i>Pinus halepensis</i>	18.59	9.5	30.62	18.4	20.51	21.49	27.94	6.6	47.13	22.31	12
<i>Pinus sylvestris</i>	34.99	28.39	23.16	15.97	27.94	3.06	29.23	14.34	51.21	25.37	13.7
Mean	19.74	28.29	25.04	24.81	21.79	17.88	22.28	17.17	62.25		
SD	14.7	18.7	4.92	13.2	5.6	13.3	10.9	12.2	22.7		

Table 4.3 Percentage (%) of current forest stands which were predicted to become unsuitable for their current species by the period 2050–2080 according to the different statistical models

Simulations using the GOTILWA+ model showed a quite stable productivity from the three species over the past century (Figure 4.5), with slight increases in production nearing the end of the century. When considering potential future climatic change with no increment in atmospheric CO₂ each of the species was predicted to reduce their production (NPP) on average by the time period 2050–2080 (Figure 4.5). This supports the results from the niche-based modeling approach, given that a reduced productivity reflects a reduction in topoclimatic suitability for these species. However, when considering the effect of increasing atmospheric CO₂ concentrations, simulated production from each of the three species showed strong increases in NPP until about 2070. After 2070, the fertilization effect of increased atmospheric CO₂ was observed to plateau, and species-specific reductions in NPP were observed. *P. halepensis* showed the strongest reduction, followed by *Q. ilex*. Although NPP rates began to decline by the end of the 21st century under the CO₂ fertilization scenario, they still maintained higher average rates (if only slightly in the case of *Q. ilex*) than those observed during the period 1950–1980.

Large differences were observed between the response of the species as modeled by GOTILWA+ and that of the multimodel ensemble. However, when considering spatially explicit simulations with a constant CO₂, the per-pixel magnitude and direction of the changes in NPP and in suitability (as predicted by the multimodel ensemble) between the period 2050–2080 and 1950–1980 were similar for two (*Q. ilex*, *P. sylvestris*) of the three species (Figure 4.6). So, less suitability for a pixel (niche-based) was reflected in less NPP for the same site, if no CO₂ effect is taken into account (process-based). On the other hand, when considering a CO₂ increment in the GOTILWA+ simulations, NPP generally showed an increase. The same per-pixel spatial trend was maintained,

where low suitability was mirrored by low NPP for *Q. ilex* and *P. sylvestris*, but the sign of the relative change in NPP vs. that of suitability, between the two focus periods, was different.

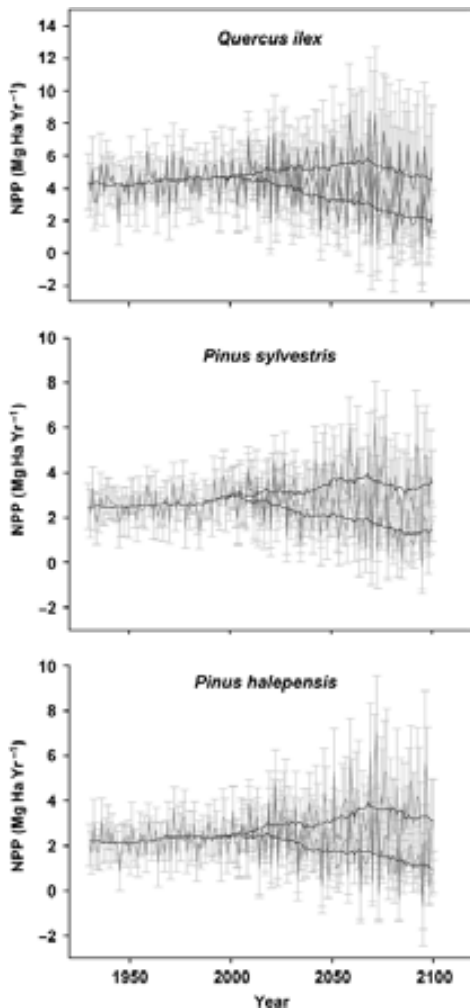


Figure 4.5 The spatially explicit change (percentage per pixel) in average per period net primary production (NPP) (GOTILWA1) and estimated Suitability (multi-niche-based model ensemble), between the periods 1950–1980 and 2050–2080, considering both GOTILWA+ simulations with (gray) and without (black) an atmospheric CO₂ increment. Lines represent linear regressions.

The root mean squared error (RMSE) between the two different modeling approaches increased by a factor of three between GOTILWA+ simulations considering atmospheric CO₂ as constant and those considering a CO₂ increment. This indicates that the introduction of CO₂ as a driver in the GOTILWA model lead to a large dissimilarity between the two modeling approaches. For GOTILWA+ simulations with a constant CO₂ concentration, the RMSE

between the percentage of change in NPP and that of suitability for the two periods was 0.22, 0.28, and 0.29 for *Q. ilex*, *P. Sylvestris*, and *P. halepensis*, respectively. The RMSE when considering a CO₂ increment was 0.64, 0.73, and 0.83 (data presented in Figure 4.6). Changes in NPP and suitability are not necessarily 1 : 1 correlated, but the RMSE between the estimates gives a measure of their similarity, and the extent of the relative dissimilarity introduced by the consideration of the potential effect of CO₂ fertilization.

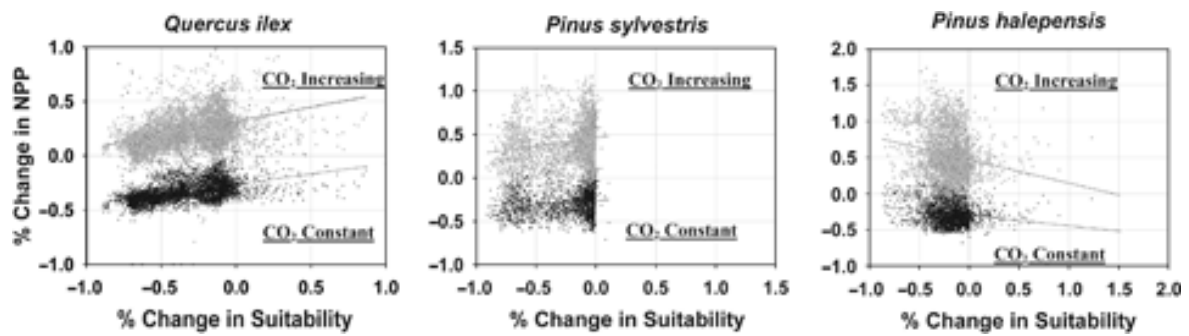


Figure 4.6 The spatially explicit change (percentage per pixel) in average per period net primary production (NPP) (GOTILWA+) and estimated Suitability (multi-niche-based model ensemble), between the periods 1950–1980 and 2050–2080, considering both GOTILWA+ simulations with (gray) and without (black) an atmospheric CO₂ increment. Lines represent linear regression

4.4 Discussion

We found that the applied niche-based models were capable of capturing the complex topoclimatic distribution of the three studied species, and that the use of a weighted multimodel ensemble improved the individual model performance. This adds to the mounting evidence that environmental conditions strongly influence species distribution patterns locally and regionally, as they do world-wide (Hawkins et al., 2003). Indeed, most of the selected variables were related to water and energy, which is consistent with the widely documented trend of plant species to be climatically driven by water–energy dynamics (see e.g., Field et al., 2005 and references therein).

An important issue regarding niche modeling is the variability of results when using different modeling techniques (Thuiller et al., 2003; Araújo et al., 2005a, b; Pearson et al., 2006). The identification of five distinct patterns of range prediction from nine models highlights the differences between modeling approaches, while providing a foundation for further investigation as to which technique, or group of techniques, may be most appropriate for predicting future ranges but inevitably calls for an ensemble forecasting to determine species distributions (Araújo & New, 2007). The best performing models are not always the same for different species, even if some of them (in particular RF, CTA, GBM) generally perform better for the species included in this study. The performance also varied according to the number of available presence records, corroborating results of other studies (Elith et al., 2006). Nevertheless, the use of different niche-based models

has been shown here to be an effective manner by which to quantify the inherent intermodel variability (Araújo et al., 2005a, b; Thuiller et al., 2005) and improve model estimates through ensemble forecast techniques. Process-based models would also benefit from such an approach, and future comparison studies should incorporate multiple process-based models.

All models, considered in any time period, entail multiple sources of uncertainty (Thuiller et al., 2003; Guisan&Thuiller, 2005). Many important biological factors are either often insufficiently described or omitted in all modeling approaches (see Guisan&Thuiller, 2005), such as small-scale environmental heterogeneity (e.g. microclimates, quantitative properties of soils), local dispersal (local dispersal leads to intraspecific aggregation; Pacala, 1997); biotic interactions across trophic levels (e.g. dispersal, pollinization; Araújo&Luoto, 2007); and processes that fragment space and create patchy aggregated distributions (e.g. forest fire events) (Fahrig, 2003). Perhaps most fundamentally for projecting possible future scenarios, large uncertainty exists regarding direct impacts of increased concentrations of atmospheric CO₂ on species physiology and competitive interactions (e.g. Ainsworth et al., 2008).

Despite their broad use, uncertainties about niche-based model predictions remain high (Hampe, 2004; Heikkinen et al., 2006; Randin et al., 2006). To date, the main drawback of niche-based models is their inability to consider important relationships such as biotic interactions, mortality, or growth (Davis et al., 1998; Hampe, 2004) and their reliance on observed distributions, which are the results of long-term historical factors (e.g., postglacial recolonization and human management), and environmental stochasticity, among other factors. As they are empirical models they are based on information relevant to present day or past species distributions. This may make their extrapolation to future scenarios questionable for some species and drivers (e.g. terrestrial vegetation and CO₂ fertilization) (Guisan&Thuiller, 2005; Pearson et al., 2006; Rickebusch et al., 2008). One technique to reduce prediction uncertainty is to fit ensembles of forecasts by simulating across more than one set of initial conditions, model classes, model parameters, and boundary conditions (see Araújo& New, 2007, for a review) and analyze the resulting range of uncertainties with probabilistic methodologies rather than using a single modeling outcome (Thuiller et al., 2006a, b; Araújo& New, 2007). Another may be to compare results from niche-based models to

those from process-based ones (e.g. Morin & Thuiller, 2009). In this study, we have shown both these techniques to be valuable in reducing and highlighting uncertainty.

The use of species level process-based models is complicated by their necessity for a large amount of data to be calibrated (often leading to the use of proxies, assumptions and expert knowledge), and large computational resources. Applications are thus restricted to well-known species for which demography or physiology has been sufficiently studied. Previous studies have highlighted large differences between different process-based model approaches (e.g., Kramer et al., 2002) and systematic difficulties in some ecosystems (Morales et al., 2005), for example, difficulties in representing soil water and soil water stress (Jung et al., 2007), or accurately predicting phenology (Keenan et al., 2009a) and related uncertainties in predicting changes to the length of the growing season. Although process-based models should theoretically be more reliable than empirical models under future climate scenarios, not all processes are fully understood (e.g., species adaptation, down-regulation, nitrogen cycling etc.), potentially also making their extrapolation to future scenarios questionable. Such uncertainties can be effectively explored through techniques such as Monte Carlo parameter estimation (e.g., Richardson et al., 2010), normally showing poorly constrained respiration processes, but well constrained canopy photosynthesis and growth. Multimodel suites, similar to that of BIOMOD, are not used for process-based models but could help reduce such uncertainties.

The effect of elevated CO₂ has been highlighted as the largest uncertainty in projecting future productivity of terrestrial vegetation (Parry et al., 2004). Elevated CO₂ stimulates photosynthetic carbon gain and net primary production over the long term despite down-regulation of Rubisco activity. It also improves nitrogen-use efficiency at both the leaf and canopy scale, while stimulating dark respiration via a transcriptional reprogramming of metabolism (Leakey et al., 2009). Experimental results indicate that plants are able to increase their water-use efficiency (WUE) as CO₂ levels rise (e.g., Picon et al., 1996; Morison, 1998), as has been corroborated under field conditions (Peñuelas & Azcón-Bieto, 1992; Ehrlinger & Cerling, 1995; Duquesnay et al., 1998; Gunderson et al., 2002; Ainsworth & Rogers, 2007). Studies have also identified interspecies variability in responses to increasing atmospheric CO₂ concentrations (e.g., Francey & Farquhar, 1982), and, importantly, have highlighted the possibility of species-specific response saturation

rates (Waterhouse et al., 2004; Betson et al., 2007). Few interspecies comparisons exist, though the general tendencies have been shown to be conserved over a large number of species (Hickler et al., 2008). It should be noted, however, that there is broad agreement that the effects of elevated CO₂ measured in experimental settings lacking potentially limiting influence of pests, weeds, nutrients, competition for resources, soil water, and air quality, may overestimate field responses of terrestrial vegetation (Long et al., 2006; Easterling et al., 2007; Tubiello et al., 2007; Ainsworth et al., 2008; Zavala et al., 2008).

Although soil water availability is the largest limitation to forest growth in Mediterranean climate regions (Allen, 2001) [and often badly represented in model projections (Hickler et al., 2009)], fertilization studies show that the availability of nutrient availability limits primary production in Mediterranean ecosystems (LeBauer&Treseder, 2008; Elser et al., 2007). Nitrogen deposition is expected to increase in Mediterranean regions in the future (Rodà et al., 2002), but nitrogen limitation is also expected to become more pronounced as atmospheric CO₂ concentration increases (the 'progressive nitrogen limitation' hypothesis) (Luo et al., 2004, 2006; de Graaff et al., 2006; Finzi et al., 2007; Reich et al., 2006). Biogeochemical models have recently incorporated dynamic nitrogen cycles (e.g., Zaehle& Friend, 2010) and results show that C–N interactions significantly reduce the stimulation of forest NPP under increased atmospheric CO₂ concentrations (e.g., Thornton et al., 2007; Jain et al., 2009). Such down-regulation in the response of forest productivity under elevated CO₂ (Ainsworth & Rogers, 2007) is estimated at about 10% for European forest species (Medlyn& Jarvis, 1999). It is thus likely that the projected future NPP (under the CO₂ enriched scenario) is overestimated in this study because it does not properly account for N down-regulation constraints (Hungate et al., 2003; Thornton et al., 2007).

Organisms are the products of chemical reactions, and their development, growth and mortality depends on various environmental factors, in particular temperature, radiation, CO₂, nutrients, and water availability. In the Mediterranean region, the future presence of a species is thus likely determined by the complex balance of temperature change, water stress. and the species-specific capacities (e.g. Peñuelas et al., 2008). Ultimately, species-specific responses may affect the structure and functioning of ecosystems (Peñuelas&Filella, 2001) due to altered competitive relationships of key performance measures and the loss of synchronization of development (Fitter

& Fitter, 2002; Gordo & Sanz, 2005). This could strongly contribute to relative fitness and thus to evolving biogeographic distributions.

The magnitude of climate change scenarios for past and future periods differ among different circulation models and therefore it is a source of uncertainty that might affect the results of the applied models (Beaumont et al., 2008; Parra & Monahan, 2008). It is therefore normally of utmost importance to apply a range of climate models and scenarios in order to estimate the inherent variability introduced by the choice of climate. In this study, due to computational limitations associated with the application of a mechanistic process-based model, we have applied only one climate scenario and model. Although the use of other climate data would change the projected distribution for each of the species, and the productivity patterns simulated by the mechanistic process-based model, we argue that the qualitative conclusions of this work are independent of the choice of climate scenario.

The presented results show that previous reports of species decline in continental Spain (e.g. Benito-Garzon et al., 2008) may be overestimated due to two reasons: the use of only one predictive niche-based model, and the failure to account for possible effects of CO₂ fertilization. Similar studies in other regions, which do not consider these two aspects, are also potentially overestimating species decline due to climate change. Similarly, the presented niche-based model results also likely overestimate the decline in suitability.

Human effects can have large impacts on the distribution of species (Chanel & Lomolino, 2000). The assumption of equilibrium between a species distribution and environmental conditions is less valid in disturbed ecosystems such as Mediterranean forests, where human influence is strong (e.g., land-use effects, fire occurrences). It has also been reported that many European tree species are not in equilibrium with climate (Svenning & Skov, 2004, not *P. sylvestris*, which was reported to be in relative equilibrium) as a consequence of postglacial dispersal limitations (Svenning et al., 2008). In this study, some of the observed imbalance between environment and spatial aggregation of tree species might be explained by the lack of equilibrium between species and current environmental conditions. It is also important to bear in mind that our results are restricted to tree species in continental Spain, and thus we can not be certain to what extent any patterns or results

that we observe here may be either affected by species occurrences in other regions, or extrapolated to other topoclimatic scenarios.

It is interesting and reassuring that changes in suitability predicted by the niche-based models conferred well with changes in NPP projected by the GOTILWA+ model (with no CO₂ increment) for two of the three studied species. This was not the case, however for *P. halepensis*. This could be explained by the fact that *P. halepensis* is distributed along the coast (Figure 4.1). Many factors other than climate can significantly influence species distributions and distribution changes (Hampe, 2004; Coudun et al., 2006; Pearson et al., 2006) and distance from the sea is used as a strong explanatory factor for presence prediction of *P. halepensis* in the niche-based models, which is not the case for the other two species. As distance from the sea is constant under climate change, this could also explain why *P. halepensis* is predicted to lose less of its current territory by the period 2050–2080 due to climate change, when compared with the other species (Table 4.3), and could in part explain the difference between projects from the niche-based models and those of GOTILWA+ for *P. halepensis*.

Ecosystems in the Mediterranean basin are prone to experience a concatenation of stochastic disturbances, including fire, drought, clearing, grazing, and land-use change. Mediterranean Basin ecosystems are thus characterized by a certain ‘unpredictability’ (Blondel & Aronson, 1999). This conditions local adaptation and manifests its effect on the phenotypic variation of forest tree species in response to macroenvironmental gradients (Volis et al., 2002). Adaptive modes could be highly important for predicting future species responses to climate change. The models presented here assume nonsignificant evolutionary and/or ecological change in a species in response to changing environmental conditions through time [thus ignoring rapid *in situ* adaptation (Thomas et al., 2001), and existing adaptation of populations to local conditions (Hampe, 2004), etc.]. Evidence suggests that species adaptation has occurred for many species (Pearman et al., 2008), implying a questionable ability of models to project species responses to potential future climates. However, we are far from a comprehensive understanding of possible species-specific adaptation capacities.

The identification of a general connection between biogeochemistry, plant physiology, disturbance, and species distributions would constitute a considerable advance in our predictive

ability (Morin et al., 2007; Chown & Gaston, 2008). Here we take the first step in using a biogeochemical model in comparison with a niche-based model, estimates of species distributions. Further work is needed to identify complementary elements of the different modeling approaches, in order to develop effective techniques for estimating species responses to potential climate change.

Plant physiology, biogeography, and related areas of research are currently merging to a new framework for understanding the patterns of the distribution of life on Earth. Ecosystem responses to climate change are driven by complex multifactor influences (Norby & Luo, 2004; Körner, 2006). An organism's niche must therefore be modeled mechanistically if we are to fully explain distribution limits (Kearney, 2006), especially when considering an organism's distribution under novel circumstances not used for the parameterization of the original model, such as a species introduction or climate change (Guisan & Thuiller, 2005). We have shown that niche-based models give accurate predictions of present species distributions (which can be improved through the use of multi model ensembles) and that comparisons with a process-based biogeochemical model can be useful in highlighting areas of uncertainty in projections under potential climate change. However, given the high variation in the accuracy of model predictions and the species-specific nature of biological responses to landscape changes (e.g. species responses to CO₂ fertilization), it seems clear that we are far from a comprehensive methodology for predicting the responses of individual species (and thus current stands) to future climatic change. Our results support recent calls for a new generation of more biologically realistic niche-based models (Guisan & Thuiller, 2005; Kearney, 2006; Araújo & Luoto, 2007; Keith et al., 2008; Rickebusch et al., 2008; Montoya et al., 2009; Nogues-Bravo, 2009). Perhaps most importantly, it is vital that models such as those used in this study are interpreted as tools for sharpening our understanding of species range constraints, and that they are only applied in a predictive capacity along with full appreciation of the uncertainty involved.

4.5 References

- Ainsworth EA, Leakey ADB, Ort DR, Long SP (2008) FACE-ing the facts: inconsistencies and interdependence among field, chamber and modeling studies of elevated [CO₂] impacts on crop yield and food supply. *New Phytologist*, **179**, 5–9.
- Ainsworth EA, Long SP (2005) What have we learned from 15 years of free-air CO₂ enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO₂. *New Phytologist*, **165**, 351–372.
- Ainsworth EA, Rogers A (2007) The response of photosynthesis and stomatal conductance to rising [CO₂]: mechanisms and environmental interactions. *Plant, Cell and Environment*, **30**, 258–270.
- Allen HD (2001) *Mediterranean Ecogeography*. Prentice-Hall, London.
- 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 MB, Cabeza M, Thuiller W, Hannah L, Williams P (2004) Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. *Global Change Biology*, **10**, 1618–1626.
- Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677–1688.
- Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, **16**, 743–753.
- Araújo MB, Miquel B, Whittaker RJ, Ladle RJ, Erhard M (2005a) Reducing uncertainty in projections of extinction risk from climate change. *Global Ecology and Biogeography*, **14**, 529–538.
- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.
- Araújo MB, Pearson RG, Thuiller W, Erhard M (2005b) Validation of species-climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.

- Bakkenes M, Alkemade J, Ihle F, Leemans R, Latour JB (2002) Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*, **8**, 390–407.
- Ball JT, Woodrow IE, Berry JA (1987) A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions. In: *Progress in Photosynthesis Research* (ed. Biggins J), pp. 221–224. Martinus-Nijhoff Publishers, Dordrecht, the Netherlands.
- Beaumont LJ, Hughes L, Pitman AJ (2008) Why is the choice of future climate scenarios for species distribution modelling important? *Ecology Letters*, **11**, 1135–1146.
- Beaumont LJ, Pitman AJ, Poulsen M, Hughes L (2007) Where will species go? Incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, **13**, 1368–1385.
- Benito-Garzon M, Sánchez de Dios R, Sainz Ollero H (2008) Effects of climate change on the distribution of Iberian tree species. *Applied Vegetation Science*, **11**, 169–178.
- Betson NR, Johannisson C, Löfvenius MO, Grip H, Granström A, Högberg P (2007) Variation in the delta¹³C of foliage of Pinussylvestris L. in relation to climate and additions of nitrogen: analysis of a 32-year chronology. *Global Change Biology*, **13**, 2317–2328.
- Blondel J, Aronson J (1999) *Biology and Wildlife of the Mediterranean Region*. Oxford University Press, Oxford, UK.
- Breiman L (2001) Random forests. *Machine Learning*, **45**, 5–32.
- Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) *Classification and Regression Trees*. Chapman & Hall, New York.
- Briceño-Elizondo R, Garcia-Gonzalo J, Peltola H, Matala J, Kellomaki S (2006) Sensitivity of growth of Scots pine, Norway spruce and silver birch to climate change and forest management in boreal conditions. *Forest Ecology and Management*, **232**, 152–167.
- Bugmann H (2001) A review of forest gap models. *Climatic Change*, **51**, 259–305.
- Busby JR (1991) BIOCLIM – a bioclimate analysis and prediction system. In: *Nature Conservation: Cost Effective Biological Surveys and Data Analysis* (eds Margules CR, Austin MP), pp. 64–68. CSIRO, Canberra.

- Campbell GS (1986) Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution. *Agricultural and Forest Meteorology*, **36**, 317–321.
- Channel R, Lomolino M (2000) Dynamic biogeography and conservation of endangered species. *Nature*, **403**, 84–86.
- Chown SL, Gaston KJ (2008) Macrophysiology for a changing world. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 1469–1478.
- Chuine I, Beaubien E (2001) Phenology is a major determinant of temperate tree range. *Ecology Letters*, **4**, 500–510.
- Coops NC, Waring RH, Schroeder T (2009) Combining a generic process-based productivity model and a statistical classification method to predict the presence and absence of tree species in the Pacific Northwest, U.S.A. *Ecological Modelling*, **220**, 1787–1796.
- Coudun C, Gegout JC, Piedallu C, Rameau JC (2006) Soil nutritional factors improve plant species distribution an illustration with *Acer campestre* (L.) in France. *Journal of Biogeography*, **33**, 1750–1763.
- Cowling RM, Rundel PW, Lamont BB, Kalin Arroyo M, Arianoutsou M (1996) Plant diversity in Mediterranean-climate regions. *Trends in Ecology and Evolution*, **11**, 362–366.
- Dai YJ, Dickenson RE, Wang YP (2004) A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance. *Journal of Climate*, **17**, 2281–2299.
- Davis AJ, Jenkinson LS, Lawton JL, Shorrocks B, Wood S (1998) Making mistakes when predicting shifts in species range in response to global warming. *Nature*, **391**, 783–786.
- Davis MB, Shaw RG (2001) Range shifts and adaptive responses to quaternary climate change. *Science*, **292**, 673–679.
- De Graaff MA, Van Groeninger KJ, Six J, Hungate BA, Van Kessel C (2006) Interactions between plant growth and nutrient dynamics under elevated CO₂: a meta analysis. *Global Change Biology*, **12**, 1–15.
- Duquesnay A, Breda N, Stievenard M, Dupouey JL (1998) Changes of tree-ring δ¹³C and water-use efficiency of beech (*Fagus sylvatica* L.) in north-eastern France during the past century. *Plant, Cell and Environment*, **21**, 565–572.

- Easterling W, Aggarwal PK, Batima P et al. (2007) Food, Fibre and Forest Products, pp. 273–313. Cambridge University Press, Cambridge, UK.
- Ehrlinger JR, Cerling TE (1995) Atmospheric CO₂ and the ratio of intercellular to ambient CO₂ levels in plants. *Tree Physiology*, **15**, 105–111.
- Elith J, Graham CH, Anderson RP et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 151.
- Elser J, Bracken M, Cleland EE et al. (2007) Global analysis of nitrogen and phosphorus limitation of primary producers in freshwater, marine and terrestrial ecosystems. *Ecology Letters*, **10**, 1135–1142.
- Fahrig L (2003) Effects of habitat fragmentation on biodiversity. *Annual Review of Ecology and Systematics*, **34**, 487–515.
- Farquhar GD, Von Caemmerer S, Berry JA (1980) A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta*, **149**, 78–90.
- Ferrier S (2002) Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? *Systems Biology*, **51**, 331–363.
- Field R, O'Brian EM, Whittaker RJ (2005) Global models for predicting woody plant richness from climate: development and evaluation. *Ecology*, **86**, 2263–2277.
- Finzi AC, Norby RJ, Calfapietra C et al. (2007) Increases in nitrogen uptake rather than nitrogen-use efficiency support higher rates of temperate forest productivity under elevated CO₂. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 14014–14019.
- Fitter AH, Fitter RSR (2002) Rapid changes in flowering time in British plants. *Science*, **296**, 1689–1691.
- Francey RJ, Farquhar GD (1982) An explanation of ¹³C/¹²C variations in tree rings. *Nature*, **297**, 28–31.
- Friedman J (1991) Multivariate adaptive regression splines. *Annals of Statistics*, **19**, 1–1141.
- Funk V, Richardson K (2002) Systematic data in biodiversity studies: use it or lose it. *Systems Biology*, **51**, 303–316.
- Gaucharel C, Guiot J, Misson L (2008) Changes of the potential distribution area of French Mediterranean forests under global warming. *Biogeosciences*, **5**, 1493–1503.

- Gibelin AL, Deque M (2003) Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. *Climate Dynamics*, **20**, 327–339.
- Giorgi F (2006) Climate change hot-spots. *Geophysical Research Letters*, **33**, L08707, doi: DOI: 10.1029/2006GL025734.
- Giorgi F, Bi X, Pal J (2004) Mean, interannual variability and trends in a regional climate change experiment over Europe. II: climate change scenarios (2071–2100). *Climate Dynamics*, **23**, 839–858.
- Gomez-Mendoza L, Arriaga L (2007) Modeling the effect of climate change on the distribution of oak and pine species of Mexico. *Conservation Biology*, **21**, 1545–1555.
- Gordo O, Sanz JJ (2005) Phenology and climate change: a long-term study in a Mediterranean locality. *Oecologia* (Berlin), **146**, 484–495.
- Gracia CA, Tello E, Sabate S, Bellot J (1999) GOTILWA: an integrated model of water dynamics and forest growth. In: *Ecology of Mediterranean Evergreen Oak Forests* (eds Rodà F, Retana J, Gracia CA, Bellot J), pp. 163–179. Springer-Verlag, Berlin.
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
- Gunderson CA, Sholtis JD, Wullschleger SD, Tissue DT, Hanson PJ, Norby RJ (2002) Environmental and stomatal control of photosynthetic enhancement in the canopy of a sweetgum (*Liquidambar styraciflua* L.) plantation during 3 years of CO₂ enrichment. *Plant, Cell and Environment*, **25**, 379–393.
- Hampe A (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography*, **13**, 469–476.
- Hastie TJ, Tibshirani R (1990) *Generalized Additive Models*. Chapman & Hall, New York.
- Hawkins BA, Field R, Cornell HV et al. (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, **84**, 3105–3117.
- Heikkinen RK, Luoto M, Araújo MB, Virkkala R, Thuiller W, Sykes MT (2006) Methods and uncertainties in bioclimatic envelope modeling under climate change. *Progress in Physical Geography*, **30**, 751–777.
- Hickler T, Fronzek S, Araújo MB, Schweiger O, Thuiller W, Sykes MT (2009) An ecosystem model-based estimate of changes in water availability differs from water proxies that are

commonly used in species distribution models. *Global Ecology and Biogeography*, **18**, 304–313.

- Hickler T, Smith B, Prentice IC, Mjofors K, Miller P, Arneth A, Sykes MT (2008) CO₂ fertilization in temperate FACE experiments not representative of boreal and tropical forests. *Global Change Biology*, **14**, 1531–1542.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Hijmans RJ, Graham CH (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272–2281.
- Hungate B, Dukes JS, Shaw MR, Luo Y, Field CB (2003) Nitrogen and climate change. *Science*, **302**, 1512–1513.
- IPCC WGI (2007) Climate Change 2007: climate change impacts, adaptation and vulnerability. In: Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007 (eds Parry ML, Canziani OF, Palutikof JP, Van Der Linden PJ, Hanson CE). Cambridge University Press, Cambridge, UK.
- Iverson LR, Prasad AM (2001) Potential changes in tree species richness and forest community types following climate change. *Ecosystems*, **4**, 186–199.
- Jain A, Yang X, Kheshgi H, McGuire AD, Post W, Kicklighter D (2009) Nitrogen attenuation of terrestrial carbon cycle response to global environmental factors. *Global Biogeochemical Cycles*, **23**, GB4028, doi: DOI: 10.1029/2009GB003519.
- Jeltsch F, Moloney KA, Schurr FM, Kochy M, Schwager M (2008) The state of plant population modelling in light of environmental change. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 171–190.
- Jung M, Le Maire G, Zaehle S et al. (2007) Assessing the ability of three land ecosystem models to simulate gross carbon uptake of forests from boreal to Mediterranean climate in Europe. *Biogeosciences*, **4**, 647–656.
- Kearney M (2006) Habitat, environment and niche: what are we modelling? *Oikos*, **115**, 186–191.
- Keenan T, Garcia R, Friend AD, Zaehle S, Gracia C, Sabate S (2009a) Improved understanding of drought controls on seasonal variation in Mediterranean forest canopy CO₂ and water

fluxes through combined in situ measurements and ecosystem modelling. *Biogeosciences*, **6**, 1423–1444.

- Keenan T, Niinemets Ü, Sabate S, Gracia C, Peñuelas J (2009b) Process based inventory of isoprenoid emissions: current knowledge, future prospects and uncertainties. *Atmospheric Chemistry and Physics*, **9**, 4053–4076.
- Keenan T, Niinemets Ü, Sabate S, Gracia C, Peñuelas J (2009c) Seasonality of monoterpene emission potentials in *Quercus ilex* and *Pinus pinea*: implications for regional VOC emissions modeling. *Journal of Geophysical Research*, **114**, D22202, doi: DOI: 10.1029/2009JD011904.
- Keenan T, Sabate S, Gracia C (2010) Soil water stress and coupled photosynthesis-conductance models: bridging the gap between conflicting reports on the relative roles of stomatal, mesophyll conductance and biochemical limitations to photosynthesis. *Agricultural and Forest Meteorology*, **150**, 443–453.
- Keenan T, Sabate S, Gracia CA (2008) Forest eco-physiological models and carbon sequestration. In: *Managing Forest Ecosystems: The Challenge of Climate Change* (eds Bravo F, LeMay V, Jandl R, Von Gadow K), pp. 83–102. Springer, Berlin.
- Keith DA, Akcakaya HR, Thuiller W et al. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560–563.
- Kharouba H, Algar A, Kerr J (2009) Historically calibrated predictions of butterfly species' range shift using global change as a pseudo-experiment. *Ecology*, **90**, 2213–2222.
- Körner C (2006) Plant CO₂ responses: an issue of definition, time and resource supply. *New Phytologist*, **172**, 393–411.
- Kramer K, Leinonen I, Bartelink HH et al. (2002) Evaluation of six process-based forest growth models using eddy-covariance measurements of CO₂ and H₂O fluxes at six forest sites in Europe. *Global Change Biology*, **8**, 213–230.
- Leakey ADB, Ainsworth EA, Bernacchi CJ, Rogers A, Long SP, Ort DR (2009) Elevated CO₂ effects on plant carbon, nitrogen, and water relations: six important lessons from FACE. *Journal of Experimental Botany*, **60**, 2859–2876.
- LeBauer DS, Treseder KK (2008) Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, **89**, 371–379.

- Leuning R, Kelliher FM, De Pury DGG, Schulze E-D (1995) Leaf nitrogen, photosynthesis, conductance and transpiration: scaling from leaves to canopies. *Plant, Cell and Environment*, **18**, 1183–1200.
- Lischke H, Zimmermann NE, Bolliger J, Rickebusch S, Löffler TJ (2006) TreeMig: a forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. *Ecological Modelling*, **199**, 409–420.
- Long SP, Ainsworth EA, Leakey ADB, Nösberger J, Ort DR (2006) Food for thought: lower-than-expected crop yield stimulation with rising CO₂ concentrations. *Science*, **312**, 1918–1921.
- Luo Y, Hui D, Zhang D (2006) Elevated CO₂ stimulates net accumulations of carbon and nitrogen in land ecosystems: a meta-analysis. *Ecology*, **87**, 53–63.
- Luo Y, Su B, Currie WS et al. (2004) Progressive nitrogen limitation of ecosystem responses to rising atmospheric carbon dioxide. *BioScience*, **54**, 731–739.
- McCullagh P, Nelder JA (1989) *Generalized Linear Models*. Chapman & Hall, New York.
- McGuire AD, Sitch S, Clein JS et al. (2001) Carbon balance of the terrestrial biosphere in the twentieth century: analyses of CO₂, climate and land use effects with four process-based ecosystem models. *Global Biogeochemical Cycles*, **15**, 183–206.
- Medlyn BE, Jarvis PG (1999) Design and use of a database of model parameters from elevated [CO₂] experiments. *Ecological Modelling*, **124**, 69–83.
- Ministerio de Agricultura PyA (2007) Tercero Inventario Forestal Nacional. Ministerio de AgriculturaPyA, Madrid, Spain.
- Montoya D, Purves DW, Urbietal IR, Zavala MA (2009) Do species distribution models explain spatial structure within tree species ranges. *Global Ecology and Biogeography*, **18**, 662–673.
- Morales P, Sykes MT, Prentice IC et al. (2005) Comparing and evaluating process-based ecosystem model predictions of carbon and water fluxes in major European forest biomes. *Global Change Biology*, **11**, 1–23.
- Morin X, Augspurger C, Chuine I (2007) Process-based modeling of tree species' distributions: what limits temperate tree species' range boundaries? *Ecology*, **88**, 2280–2291.

- 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.
- Morison JIL (1998) Stomatal response to increased CO₂ concentration. *Journal of Experimental Botany*, **49**, 443–452.
- New M, Hulme M, Jones PD (1999) Representing twentieth century space-time climate variability. Part 1: development of a 1961–1990 mean monthly terrestrial climatology. *Journal of Climate*, **12**, 829–856.
- Nigh GD, Ying CC, Qian H (2004) Climate and productivity of major conifer species in the interior of British Columbia Canada. *Forest Science*, **50**, 659–671.
- Ninyerola M, Pons X, Roure JM (2007a) Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. *International Journal of Climatology*, **27**, 1231–1242.
- Ninyerola M, Pons X, Roure JM (2007b) Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. *Theoretical and Applied Climatology*, **89**, 195–209.
- Nogues-Bravo D (2009) Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, **18**, 521–531.
- Norby RJ, Luo Y (2004) Evaluating ecosystem responses to rising atmospheric CO₂ and global warming in a multi-factor world. *New Phytologist*, **162**, 281–293.
- Pacala SW (1997) Dynamics of plant communities. In: *Plant Ecology* (ed. CrawleyMJ), pp. 532–555. Blackwell Scientific Publications, Oxford.
- Pacala SW, Canham CD, Silander JA Jr (1993) Forest models defined by field measurements: 1. The design of a northeastern forest simulator. *Canadian Journal of Forest Research*, **23**, 1980–1988.
- Parra JL, Monahan WB (2008) Variability in 20th century climate change reconstructions and its consequences for predicting geographic responses of California mammals. *Global Change Biology*, **14**, 2215–2231.
- Parry ML, Rosenzweig C, Iglesias A, Livermore M, Fischer G (2004) Effects of climate change on global food production under SRES emissions and socio-economic scenarios. *Global Environmental Change-Human and Policy Dimensions*, **14**, 53–67.

- Pearman PB, Guisan A, Breonmann O, Randin CF (2008) Niche dynamics in space and time. *Trends in Ecology and Evolution*, **23**, 149–158.
- Pearson RG, Thuiller W, Araújo MB et al. (2006) Model-based uncertainty in species' range prediction. *Journal of Biogeography*, **33**, 1704–1711.
- Pelkonen P, Hari P (1980) The dependence of the springtime recovery of CO₂ uptake in Scots pine on temperature and internal factors. *Flora*, **169**, 398–404.
- Peñuelas J, Azcón-Bieto J (1992) Changes in leaf $\Delta^{13}\text{C}$ of herbarium plant species during the last 3 centuries of CO₂ increase. *Plant, Cell and Environment*, **5**, 485–489.
- Peñuelas J, Filella I (2001) Phenology – responses to a warming world. *Science*, **294**, 793–795.
- Peñuelas J, Hunt JM, Ogaya R, Jump AS (2008) Twentieth century changes of tree-ring $\delta^{13}\text{C}$ at the southern range-edge of *Fagus sylvatica*: increasing water-use efficiency does not avoid the growth decline induced by warming at low altitudes. *Global Change Biology*, **14**, 1076–1088.
- Peterson AT (2007) Uses and requirements of ecological niche models and related distributional models. *Biodiversity Informatics*, **3**, 59–72.
- Picon C, Guehl JM, Aussenac G (1996) Growth dynamics, transpiration and water-use efficiency in *Quercus robur* plants submitted to elevated CO₂ and drought. *Annals of Forest Science*, **53**, 431–446.
- Randin CF, Dirnböck T, Dullinger S, Zimmermann NE, Zappa M, Guisan A (2006) Are niche-based species distribution models transferable in space? *Journal of Biogeography*, **33**, 1689–1703.
- Reich PB, Hobbie SE, Lee T et al. (2006) Nitrogen limitation constrains sustainability of ecosystem response to CO₂. *Nature*, **440**, 922–925.
- Richardson AD, Williams M, Hollinger D et al. (2010) Estimating parameters of a forest ecosystem C model with measurements of stocks and fluxes as joint constraints. *Oecologia* (Berlin), doi: DOI: 10.1111/j.1466-8238.2009.00478.x.
- Rickebusch S, Thuiller W, Hickler T, Araújo MB, Sykes MT, Schweiger O, Lafourcade B (2008) Incorporating the effects of changes in vegetation functioning and CO₂ on water availability in plant habitat models. *Biology Letters*, **4**, 556–559.

- Ricklefs R (2004) A comprehensive framework for global patterns in biodiversity. *Ecological Letters*, **7**, 1–15.
- Ridgeway G (1999) The state of boosting. *Computational and Scientific Statistics*, **31**, 172–181.
- Ripley BD (1996) Pattern Recognition and Neural Networks. Cambridge University Press, Cambridge, UK.
- Rodà F, Avila A, Rodrigo A (2002) Nitrogen deposition in Mediterranean forests. *Environmental Pollution*, **118**, 205–213.
- Rushton SP, Ormerod SJ, Kerby G (2004) New paradigms for modelling species distributions? *Journal of Applied Ecology*, **41**, 193–200.
- Scheller RM, Mladenoff DJ (2007) An ecological classification of forest landscape simulation models: tools and strategies for understanding broad-scale forested ecosystems. *Landscape Ecology*, **22**, 491–505.
- Schröter D, Cramer W, Leemans R et al. (2005) Ecosystem service supply and vulnerability to global change in Europe. *Science*, **25/310**, 1333–1337.
- Sitch S, Huntingford C, Gedney N et al. (2008) Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs). *Global Change Biology*, **14**, 2015–2039.
- Skov F, Svenning JC (2004) Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography*, **27**, 366–380.
- St Clair JB, Howe GT (2007) Genetic maladaptation of coastal Douglas-fir seedlings to future climates. *Global Change Biology*, **13**, 1441–1454.
- Svenning JC, Normand S, Skov F (2008) Postglacial dispersal limitation of widespread forest plant species in nemoral Europe. *Ecography*, **31**, 316–326.
- Svenning JC, Skov F (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, **7**, 565–573.
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285–1293.
- Thomas CD, Bodsworth EJ, Wilson RJ, Simmons AD, Davies ZG, Musche M, Conradt L (2001) Ecological and evolutionary processes at expanding range margins. *Nature*, **411**, 577–581.

- Thomas CD, Cameron A, Green RE et al. (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thornton PE, Lamarque J-F, Rosenbloom NA, Mahowald NM (2007) Influence of carbon–nitrogen cycle coupling on land model response to CO₂ fertilization and climate variability. *Global Biogeochemical Cycles*, **21**, doi: DOI: 10.1029/2006GB002868.
- 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 (2007) Biodiversity-climate change and the ecologist. *Nature*, **448**, 550–552.
- Thuiller W (2010) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020–2027.
- Thuiller W, Albert CH, Araujo M et al. (2008) Predicting global change impacts on plant species distributions: future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 137–152.
- Thuiller W, Araujo MB, Lavorel S (2003) Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, **14**, 669–680.
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD – a platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.
- Thuiller W, Lavorel S, Araujo MB, Sykes MT, Prentice C (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences*, **102**, 8245–8250.
- Thuiller W, Lavorel S, Sykes MT, Araújo MB (2006a) Using niche-based modelling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distributions*, **12**, 49–60.
- Thuiller W, Midgley GF, Hughes GO, Bomhard B, Drew G, Rutherford MC, Woodward FI (2006b) Endemic species and ecosystem vulnerability to climate change in Namibia. *Global Change Biology*, **12**, 759–776.
- Tubiello FN, Soussana JF, Howden SM (2007) Crop and pasture response to climate change. *Proceedings of the National Academy of Sciences*, **104**, 19686–19690.
- Underwood EC, Viers JH, Klausmeyer KR, Cox RL, Shaw MR (2009) Threats and biodiversity in the Mediterranean biome. *Diversity and Distributions*, **15**, 188–197.

- Van Der Meer PJ, Jorritsma ITM, Kramer JK (2002) Assessing climate change effects on long-term forest development: adjusting growth, phenology and seed production in a gap model. *Forest Ecology and Management*, **162**, 39–52.
- Volis S, Mendlinger S, Turuspekov Y, Esnazarov U (2002) Phenotypic and allozyme variation in Mediterranean and desert populations of wild barley, *Hordeum spontaneum*. *Evolution*, **56**, 1403–1415.
- Wang Y-P, Leuning R (1998) A two-leaf model for canopy conductance, photosynthesis and partitioning of available energy I: model description and comparison with a multi-layered model. *Agricultural and Forest Meteorology*, **91**, 89–111.
- Waterhouse JS, Switsura VR, Barkera AC, Cartera AHC, Hemming DL, Loaderd NJ, Robertson I (2004) Northern European trees show a progressively diminishing response to increasing atmospheric carbon dioxide concentrations. *Quaternary Science Reviews*, **23**, 803–810.
- Zaehle S, Friedlingstein P, Friend AD (2010) Terrestrial nitrogen feedbacks may accelerate future climate change. *Geophysical Research Letters*, **37**, L01401, doi: DOI: 10.1029/2009GL041345.
- Zaehle S, Friend A (2010) Carbon and nitrogen cycle dynamics in the O-CN land surface model, I: model description, site-scale evaluation and sensitivity to parameter estimates. *Global Biogeochemical Cycles*, doi: DOI: 10.1029/2009GB003521.
- Zavala JA, Casteel CL, DeLucia EH, Berenbaum MR (2008) Anthropogenic increase in carbon dioxide compromises plant defense against invasive insects. *Proceedings of the National Academy of Sciences*, **105**, 5129–5133.
- Zierl B, Bugmann H (2007) Sensitivity of carbon cycling in the European Alps to changes of climate and land cover. *Climatic Change*, **85**, 195–212.

5. ON PROJECTIONS OF MULTIPLE SPECIES: AGGREGATED VERSUS INDIVIDUAL PLANT SPECIES DISTRIBUTIONS

STUDY: ‘Coexistence of *Abies alba* (Mill.) – *Fagus sylvatica* (L.) and climate change impact in the Iberian Peninsula: A climatic-niche perspective approach’

5.1 Introduction

The effects of global climate change on vegetation and forests might induce deep transformations in natural resources and landscape structure, posing a challenge to biodiversity and habitats (Thomas et al., 2004; Botkin et al., 2007; Nogués-Bravo et al., 2007). The implications of climate warming for vegetation have therefore been widely studied on different scales, ranging from the physiological responses of plants (Peñuelas et al., 2002; Walther et al., 2002) to regional plant distribution (Thuiller et al., 2005), using various modeling approaches (Guisan & Zimmermann, 2000; Elith et al., 2006).

Species distribution models (SDM) refer to the whole set of statistical correlative approaches that extrapolate the environmental data associated with a species’ presence and/or absence and project its relationship with these data onto a different sites and/or time period. As they are easy to implement, these models are now widely used to address various issues in ecological research, while also providing guidance for applied research (Franklin, 2010).

The theoretical basis for most species distribution models is the niche theory, which establishes the existence of a whole set of conditions and resources in which a species can live (sensu Hutchinson, 1957; see Holt, 2009). The niche theory makes it possible to use current patterns of distribution to characterize the realized niche (RN), whereas the fundamental niche (FN) concept indicates the whole range of conditions in which a species may survive and reproduce.

There has been great debate on the scientific scene over what these species distribution models

are really modeling (Hirzel & Lay, 2008; Kearney, 2006) and many studies point out that the proximity of a modeled niche to either a realized or a fundamental one is dependent on the techniques used, the selection of absences and whether the model's construction is built on presence-absence data or presence data alone (Jiménez-Valverde et al., 2008). Furthermore, it has often been argued that the processes involved (e.g., growth, reproduction, competition and migration) are not evaluated directly, thus hampering both the interpretation and application of these models (see Guisan & Thuiller, 2005; Thuiller et al., 2008, for an extended review).

These models have usually treated species distribution on an individual basis, although vegetation units have also been considered (Miller & Franklin, 2002). Community assemblies may be modeled via a number of different strategies: (1) assemble first, predict species later; (2) predict species first, assemble later; or (3) assemble and predict at the same time (Ferrier and Guisan, 2006). Modeling communities (several species simultaneously) may provide a faster way to model diversity and examine different patterns of co-occurrence. Although community-based models may render accurate results, species may respond differently to variations in climate (Williams & Jackson, 2007) and are subject to differences in plasticity (Valladares et al., 2007), prejudicing these models' ability to predict various situations in space and time (Guisan and Zimmermann, 2000) and therefore be extrapolated (Elith and Leathwick, 2009). Previous studies investigated the effect of individual models versus community models (Baselga & Araújo, 2009) using a community approach based on patterns of co-occurrence and co-exclusions (Canonical Quadratic Ordination). Their results pointed out some shortcomings due to the generality of the model and species interactions in such community models.

In fact, by applying SDMs to model community distribution, we would expect the projection of the current conditions in which a community appears (realized niche of the community; RNC) to result in smaller areas than the overlap of the projections of single species. If this is true, models of future communities may overestimate species assemblages if produced by Boolean intersection (overlap) from the performance of individual species. Such an overlap could be interpreted as the niche space lying between the realized niche of the community and the fundamental niche, corresponding to the pseudo-fundamental niche of a community (PseudoFNC; Figure 5.1). Although the fundamental niche of any species or community is difficult to estimate, the pseudo-

fundamental niche represents combinations of the environmental space in which the species in a community currently occur, projecting a niche close to the fundamental niche. Therefore, the degree of disagreement between both niches (realized to pseudo-fundamental) may elucidate whether a community's current realized niche meets the whole set of current conditions met by species comprising the target community ($\text{Realized Niche sp1} \cap \text{Realized Niche sp2}$).

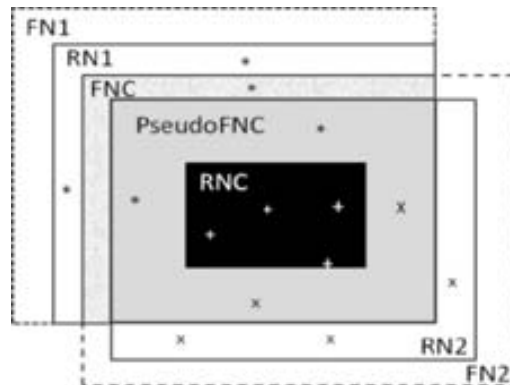


Figure 5.1 Framework of the niche concept. Fundamental niche of the community (FNC) corresponds to the intersection of the Fundamental niche of indicator species of the community (FN1, FN2). Pseudo-fundamental niche of the community (PseudoFNC) models the intersection of the realized niches of each indicator species of the community (RN1, RN2; modeled by the overlap of individual models) whereas the realized niche of the community (RNC) is modeled by the co-occurrence of the indicator species of the community.

In the Iberian Peninsula, the distributions of silver fir (*Abies alba* Mill.) and European beech (*Fagus sylvatica* L.) and *Abieti-Fagetum* community are restricted to some areas of the Eurosiberian bioregion (Figure 5.2). Therefore, here both species are located on the southwest boundary of its European distribution, making them more vulnerable to changes in climate and management (Jump et al., 2006; Aitken et al., 2008). Although both species show similar distribution patterns within this region, the community *Abieti-Fagetum* is relatively scarce and the two indicator species often occur separately. Fir is mainly distributed at higher altitudes than beech, which requires greater humidity and shows a more continental distribution in the European context (European Flora Atlas: Lahti & Lampinen, 1999).

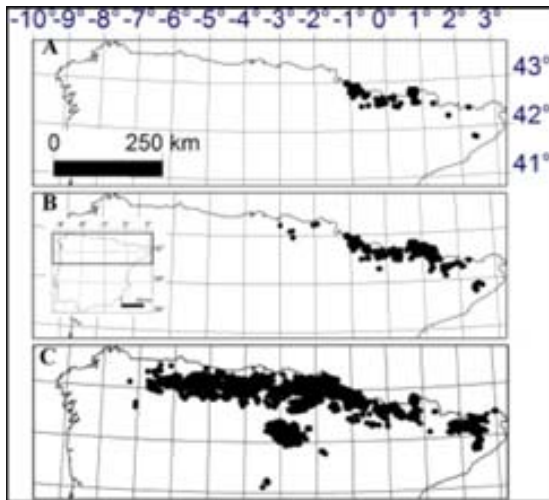


Figure 5.2 Spanish national forest inventory plot location of the co-occurrence of *Abie salba* and *Fagus sylvatica* (A), of *Abies alba* (B) and of *Fagus sylvatica* (C) in the Pyrenees and in the Iberian Peninsula.

Climate change scenarios (IPCC, 2007) predict a rise in temperature and changing patterns of precipitation in the Iberian Peninsula, resulting in increased water deficit. The relatively high water requirements of these species therefore will reduce the viability of the community of *Abies alba* and *Fagus sylvatica*. All these elements make the case of the *Abies alba*- *Fagus sylvatica* a good study-system for analyzing the performance of species-

based distribution models when building community assembly models. Previous studies already presented community model results for the *Abieti-Fagetum* community in Switzerland (Brzeziecki et al., 1993, 1995), but did not directly consider different modeling approaches.

In the present research we aim to assess the future distribution and the reliability of predictions about the community formed by the co-occurrence of *Abies alba* and *Fagus sylvatica* in the Iberian Peninsula, by considering and comparing both the individual-species approach (OIM, strategy 2) and the community-based approach (CM, strategy 1), under climate change scenarios A1FI and A2 (IPCC, 2007).

5.2 Material and methods

We used a Generalized Linear Model (GLM) to map the suitability of *Abies alba*, *Fagus sylvatica* and mixed forests of these species based on presence/absence plots from the third National Forest Inventory (Ministerio de Agricultura PyA, 2007). The National Forest Inventory uses a regular sampling of all 1×1 UTM grid intersections matching forest areas. A presence of the target-species was selected only when dominant (first or second more abundant species according to their Basal Area [BA]).

Data on absences may influence model accuracy (Chefaoui and Lobo, 2008; Lobo et al., 2010), so we built 250 different data sets and run models for each dataset to obtain robust results, i.e., results not dependent on absences. Prevalence was kept constant ($N_{\text{absence}} = N_{\text{presence}}$) in each dataset and we imposed a distance constraint on absence selection: plots within a buffer zone of 5km around presences were not considered. By imposing these criteria we restricted the selection of absences in suitable climatic conditions.

We followed two approaches to project community suitability: (1) overlapping individual-species modeling (OIM), and (2) community-based modeling (CM) of the co-occurrence of both species on the same forest inventory stand.

Selection of variables

Firstly, a correlation analysis was conducted between several climatic and topographic variables that explain, a priori, the distribution of *Abies alba* and *Fagus sylvatica* (Benayas et al., 2002; Thuiller et al., 2003): minimum, maximum and mean temperature and precipitation on a seasonal and yearly time scale, and water availability, computed as precipitation minus evapotranspiration (Thornthwaite, 1948).

Climatic variables were derived from the Digital Climatic Atlas of the Iberian Peninsula (Ninyerola et al., 2000; available at: <http://opengis.uab.es/wms/iberia/en/index.htm>). This cartography consists of 65 monthly maps of mean air temperature (minimum, mean and maximum), precipitation and solar radiation derived from 1000 to 2000 meteorological stations, depending on the variable. These climatic grids are built using multiple regression and interpolation techniques, such as inverse distance weighting and kriging, coupled with a residual error correction method implemented in a GIS environment. Data from meteorological stations have been combined with altitude, latitude, distance from coast, solar radiation and terrain curvature to obtain 180 m spatial resolution grids of every climatic variable. Cross validation results show a root mean square error (Root Mean Square Error, RMSE) of 6–20 mm for annual precipitation, and 0.8–1.5°C for annual mean temperatures. Topographic variables such as slope, terrain curvature, solar radiation and continentality (cost-

distance to coast) were derived from a 200m spatial resolution Digital Elevation Model (DEM).

Highly correlated variables were removed for subsequent species distribution modeling. We preferred this procedure to the use of factors extracted from Principal Components Analysis, as when many variables are used it is often difficult to interpret the extracted components in biological terms. Correlation analysis was performed using R 2.11.1 software (R Development Core Team, 2010) and correlations above 0.70 led to the elimination of one of the correlated variables. To decide which variables should be selected, we used the rule of the most comprehensive and integrative variable. For example, water balance is obviously highly correlated with precipitation and temperature but since it is also the most integrative variable, we would then remove precipitation and temperature from further analysis. The variables used in each model are shown in Table 5.1.

MODELING APPROACH	VARIABLES	MEAN EXPLAINED VARIANCE	AUC
IM <i>Abies alba</i> (individual model)	WAsummer; WAwinter MinWiT; Curvature; Slope; RAsummer	74 %	0.98
IM <i>Fagus sylvatica</i> (individual model)	WAspring; MinWiT Friction; Slope; RAsummer	65%	0.96
CM <i>Fagus sylvatica</i> + <i>Abies alba</i> (community model)	WFall; MinWiT; Curvature; Slope	84%	0.98

Table 5.1 Summary of the GLMs. Variables: WA (Water Availability), MinWiT (minimum winter temperature), CURV (Terrain curvature), SLP (Slope), RA (Solar radiation). AUC refers to the Area Under the Receiver Operating Characteristics curve (ROC), a measure of accuracy calculated with the 20% of points in each dataset. Data correspond to the mean values of 250 datasets with different absence locations.

Generalized linear models for *Abies alba*, *Fagus sylvatica* and *A. alba*-*F. sylvatica* co-occurrence

A stepwise GLM was performed using R software (R Development Core Team, 2010) based on an Akaike Information Criterion (AIC; Akaike, 1974). This procedure was run separately for *Abies alba*, *Fagus sylvatica* and the community *Abies alba*-*Fagus sylvatica*, with presence and absence considered on a scale of 1 km². In each approach 250 iterations (runs) of a GLM were performed, using different selections of absences. For each run, the GLM algorithm chose the best combination of variables that minimize the input information in the model (AIC), and the most repeated set of variables in the 250 models was chosen to run the final model. The final model consisted of a regression using averaged regression coefficients from the 250 GLM iterations, using the selected formula.

Model calibration was performed by using 80% of the plots from each dataset (250 in total), with 20% of the plots reserved to calculate the accuracy of the model using the area under the Receiver Operating Characteristics curve parameter (Fielding & Bell, 1997). The final model accuracy for each approach was computed using the mean of the 250 runs.

Suitability models for each species and CM produced an output (suitability index), ranging from 0 to 1, which can be mapped. A threshold needs to be set in order to differentiate a suitable location from unsuitable ones (binary response). We calculated this threshold as a compromise between maximizing the correct predicted presences and minimizing unpredicted presences, thus optimizing model sensitivity and specificity (see examples in Pearce & Ferrier, 2000; Randin et al., 2009).

Future climatic scenarios: downscaling global circulation models

The IPCC socioeconomic analysis (IPCC, 2007) proposes several scenarios of CO₂ emissions. Different scenarios and climate variability may achieve quite different results (Beaumont et al., 2007), so we used two climatic projections (A1FI, A2) derived from the coupled atmosphere-ocean general circulation model (AOGCM) HadCM3 simulation, developed at the Hadley Centre – UK

(Gordon et al., 2000; Pope et al., 2000). Each scenario represents a storyline in greenhouse gases emissions, taking into account demographic and social frameworks. In our study we selected two scenarios from the A1 and A2 storylines family. The former forecasts a future of rapid economic growth and intensive fuel use (A1), whereas the latter describes growth, albeit limited to a regional level and free of any rapid technological development (IPCC, 2007). In the Iberian Peninsula these scenarios describe varying degrees of severity: A1FI predicts an increase of 4.5°C in mean annual temperature and a drop of 111 mm in annual precipitation, whereas A2 represents a milder change, with an increase of 3.2°C and an average decrease in annual precipitation of 95 mm.

These general circulation models were applied to obtain the mean annual temperature and precipitation for the period 2050–2080 on a grid with a 4000 m spatial resolution. We downscaled these values by adding the predicted mean increase or decrease in temperature and precipitation (2050–2080) to the Digital Climatic Atlas of the Iberian Peninsula, which offers a finer resolution for mountainous areas (Randin et al., 2009). The topographic variables remained constant for the two scenarios. The new values for the models' variables resulted in new suitability maps in the predicted scenarios, making it possible to determine future increases or decreases in suitability.

5.3 Results

Individual-species models present positive evaluation results with AUC values of 0.98 for *Abies alba* and 0.96 for *Fagus sylvatica* (Table 5.1). The geographic pattern of the topoclimatic area suitable for *Abies alba* is restricted to the northern mountain regions of the Iberian Peninsula, although some remnants of suitable areas may also be found in the mountainous central areas of Spain (Figure 5.3A). Curiously, large suitable areas can be found in the northeastern mountains, where the species is not currently present, although it has been reported to exist there during the Quaternary (Peñalba, 1994; Benito Garzón et al., 2007).

With respect to the present time, the models predict a larger suitable area for *Fagus sylvatica* compared to *Abies alba* (Table 5.2). Even though the range of *Fagus sylvatica* is narrow, but larger than *Abies alba*, the variability of climatic situations within the current presences provide suitable

areas in other locations, such as the central mountain areas of Spain, the mountain regions in the east of the Iberian Peninsula and even the southern mountain region (Figure 5.3D).

SPECIES	Suitable area (km ²)			Change in suitable area	
	PRESENT	A2	A1FI	PRESENT-A1FI	PRESENT-A2
<i>Abies alba</i>	19418	8625	7735	-60 %	-56 %
<i>Fagus sylvatica</i>	56024	9287	6241	-88 %	-83 %

Table 5.2 Total suitable area of *Abies alba* and *Fagus sylvatica*

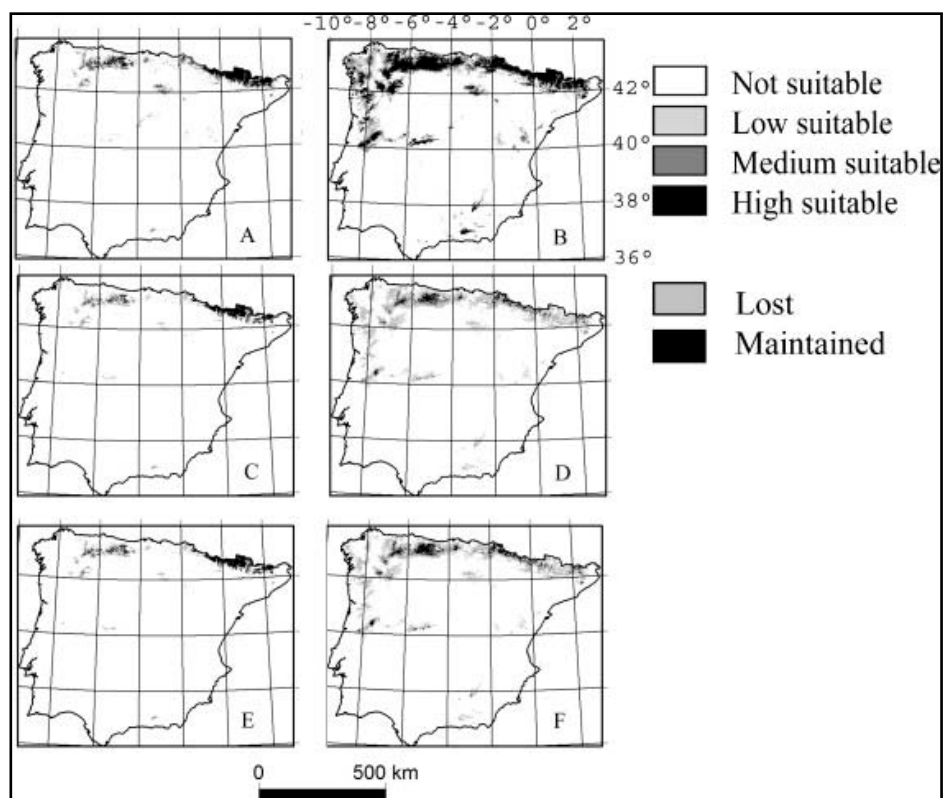


Figure 5.3 Current and future suitability for *Abies alba* and *Fagus sylvatica* separately. Present suitability index for *Abies alba* (A), predicted dynamics of suitable areas in scenario A1FI (C) and A2 (E). Present suitability index for *Fagus sylvatica* (B), predicted dynamics of suitable areas in scenario A1FI (D) and A2 (F).

Future scenarios predict a large decrease for both species. This decrease is stronger for *Fagus sylvatica* than for *Abies alba* but the suitable area for *Abies alba* remains smaller in both the climate change scenarios (see Figure 5.3B, E and F and Table 5.2). Suitable area decrease is linked to the aridity degree assumed by the scenario and, as expected, a larger loss of suitable area is predicted for both species under the A1FI scenario than under then A2 scenario. No gain in suitable area is observed in any of the scenarios considered.

The geographic pattern of suitable area loss is similar for both species: the western Iberian peninsula and forests at lower altitudes are much more affected while the Pyrenean range (North-East) remains stable, with the exception of the west side, where losses are forecasted to be large (Figure 5.3B–E).

The Community Model (CM) also showed good accuracy results (Table 5.1) and correctly identified the 77.2% of the stands in which both species are present. A ROC curve cannot be calculated for the overlap of individual models (OIM) approach since it is a simple GIS overlap function, although it correctly classified the 88.2% of co-occurrences (Table 5.5).

As expected, the suitable area for mixed forests obtained by overlapping individual (OIM) models is larger than that obtained by the community models approach (CM), regardless of the climate change scenario considered (Table 5.3). Both the CM and the OIM predict that the co-occurrence of these two species will undergo a large decrease insuitable area but the percentage of area reduction is higher in the OIM than in the CM (Table 5.3). Of the climate change scenarios, the CM is more sensitive than the OIM, predicting a larger reduction of suitable area in the A2 scenario. The CM suitable area represents 70% of the OIM area suitable for current conditions, and this figure drops to 50% and 40% in the A1FI and A2 scenarios, respectively (Table 5.3).

Model Approach	Suitable area (km ²)			Change in suitable area	
	PRESENT	A1FI	A2	PRESENT-A1FI	PRESENT-A2
OIM	16105	1787	2722	-91.8 %	-90.2 %
CM	11238	927	1096	-88.9 %	-83.1 %

Table 5.3 Differences in suitable area and predicted changes for each modeling approach: OIM (Overlapped Individual Models) and CM (Community model).

Geographically, both approaches project the most suitable area for the community in the Pyrenees under present conditions (Figure 5.4). Despite this common tendency, there are differences between the two approaches as regards the spatial distribution of the suitable area: the overlap of individual species models forecasts the occurrence of mixed forests in the Cantabrian mountain range system (northwest Iberian Peninsula), whereas the community model locates small patches of suitable area along the southern face of the Pyrenees (Figure 5.4).

In both the climate change scenarios considered, the Cantabrian Mountains and inland mountain ranges are sites where the decrease in suitable area is predicted to be more severe, whereas it will remain fairly stable in the Pyrenees, especially in places at higher altitudes (Figure 5.4, A1FI and A2).

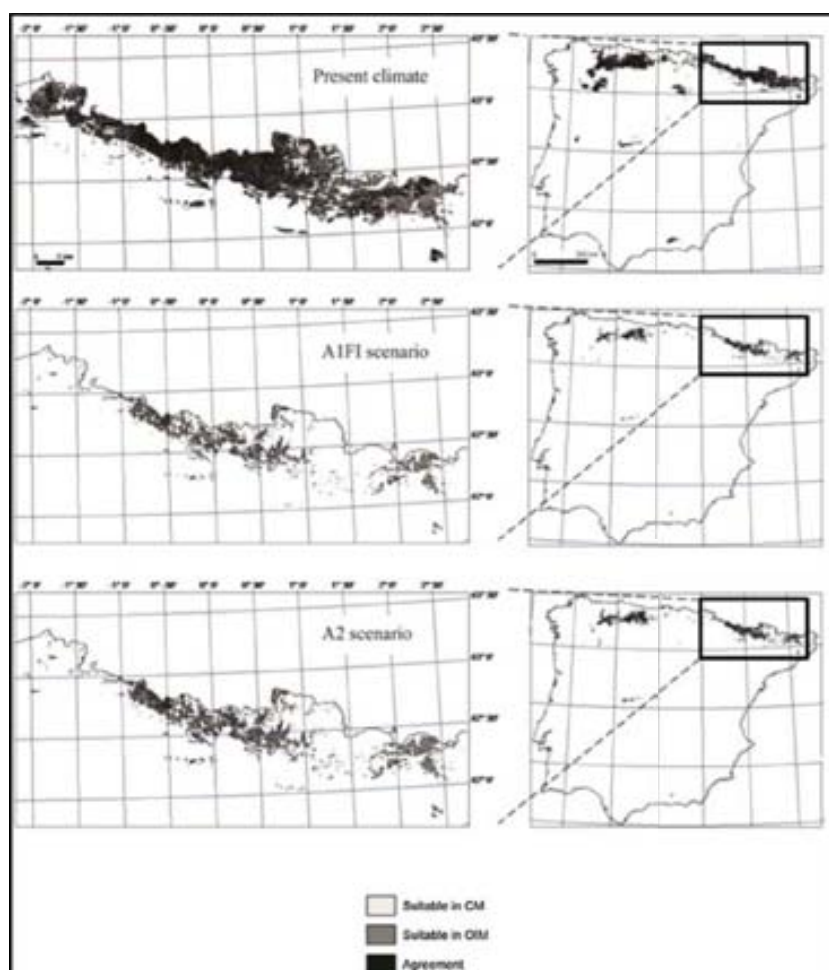


Figure 5.4 Agreement and disagreement between modeling Community and Overlapped Individual Models approach (CM and OIM) for present climate and future scenarios. Agreement declines with climate change and most suitable area is only forecasted by OIM approach. Main agreement geographical region corresponds to the central Pyrenees range.

The degree of spatial disagreement, calculated as the area predicted to be suitable by only one of the approaches, increases when the models are projected onto climate change scenarios (Table 5.4). As expected, the realized niche of the community (RNC; depicted from the community model) is geographically included in the Pseudo-fundamental niche of the community (PFNC; depicted from the overlap of individual models) for the current environmental conditions, i.e. the suitable area predicted only by the overlap of individual models is 30% compared to 0.6% in the community model (Table 5.4).

	PRESENT (km ²)	A1FI(km ²)	A2(km ²)
Suitability agreement (OIM and CM)	11 170	464	1 030
Disagreement OIM (suitable only in OIM)	4 935 (30%)	1 323 (74%)	1 692 (62%)
Disagreement CM (suitable only in CM)	68 (0.6%)	463 (50%)	66 (6%)

Table 5.4 Agreement between the community model (CM) and overlay individual model (OIM) for the climate change scenarios considered. Percentages in brackets indicate the degree of disagreement in suitable areas (Area of disagreement under approach X/Total suitable area under approach X).

This geographical inclusion of the modeled niche also occurs in the A2 scenario, with a 6% disagreement in the community model, although the spatial disagreement increases in both approaches for both climate change scenarios. In the case of the most severe scenario, A1FI, the geographical inclusion effect of the community model in the overlap individual model disappears because spatial disagreement rises up to 50%.

The spatial differences between the modeling approaches become more relevant when assessing the future of the current plots of mixed forests of *Abies alba* and *Fagus sylvatica*. The overlap of individual models (OIM) predicts a loss of between 73% and 80% of the current stands

depending on the scenario, whereas the community model (CM) increases this loss to between 92% and nearly 99% of the current co-occurrences of the forest inventory for these species (Table 5.5). In many cases, the OIM predicts that losses of mixed forests are the result of a decline in *Fagus sylvatica*, while *Abies alba* would remain on these sites (Table 5.5). Remnants of future suitable plots are to be found on the northern slopes of Pyrenees and at higher altitudes (Table 5.6), which suggests an upwards shift in its lower altitudinal limit. Likewise, more threatened forests are those located at lower altitudes.

	%	NOT SUITABLE FOR MIXED FORESTS	SUITABLE FOR <i>Fagus sylvatica</i>	NOT SUITABLE FOR ANY SPECIES	SUITABLE FOR <i>Abies alba</i>
OIM	PRESENT	15.8	7.0	0.6	8.2
CM		22.8		22.8	
OIM	A1FI	80.3	19.7	14.6	45.9
CM		98.7		98.7	
OIM	A2	73.9	24.2	8.3	41.4
CM		92.4		92.4	

Table 5.5 Percentage of predicted suitable/non-suitable topo-climatic conditions for current NFI plots of mixed forest (*Abies alba*-*Fagus sylvatica*) for the two model approaches (OIM: overlap individual model; CM: community model) under A1FI and A2 scenarios. Not suitable areas in the case of OIM are divided, depending on the source of the loss of the co-occurrence. Shaded cells correspond to present climate plots (model accuracy in OIM).

5.4 Discussion

The present study highlights a decline in the climatic suitability of *Fagus sylvatica* and *Abies alba* and its present co-occurrence in the Iberian Peninsula. This trend can be appreciated in both the future climatic scenarios and the two modeling approaches for co-occurrence assembly (individual overlapping and community).

This pattern concurs with results of modeling studies on both a European (Bakkenes et al., 2002; Thuiller et al., 2006) and an Iberian Peninsula scale (Benito Garzón et al., 2008). Furthermore, indications of this downward trend are already being detected in fir and beech forests in the region. For instance, Macias et al. (2006) documented a die-back of Pyrenean forests of *Abies alba* in recent decades, associated with the interaction of climate tendencies and management history. Likewise, Puddu et al. (2003) detected vulnerability of *Abies alba* to pathogenic fungi in less suitable climatic conditions, while canopy defoliation of *Fagus sylvatica* forests on a regional scale was detected by remote sensing during a drought episode that affected Western Europe in 2003 (Lloret et al., 2007). Xeric conditions have also been reported as affecting the physiological performance of *Abies alba* (Peguero-Pina et al., 2007). These episodes, and overall water stress, are likely to increase with climate change (Meehl & Tebaldi, 2004; Schar & Jendritzky, 2004), reducing the regional climatic fitness of the species. In keeping with our predictions, an upward shift in the altitudinal distribution of *Fagus sylvatica* has been detected in the region (Peñuelas and Boada, 2003; Jump et al., 2006), similarly to the cases reported for other species and regions (Parmesan & Yohe, 2003; Beckage et al., 2008; Kelly & Goulden, 2008; Lloret & González-Mancebo, 2011), including latitudinal changes (Jump et al., 2009). The expected upward shift for the two species studied predicts a possible competition between *Abies alba* and *Pinus uncinata*. According to Ameztegui and Coll (2011) such chorological changes in the Pyrenees may take place since silver fir may find suitable conditions for colonizing pine dominated stands.

Both approaches to co-occurrence modeling (CM and OIM) indicate that the Pyrenees will be the most stable suitable area for the maintenance of these mixed forests. For this reason, it is important to consider the Pyrenees as a future potential refuge for these species and a reservoir for mixed silver fir-beech forests in the Iberian Peninsula. Interestingly, this mountain range has been a climate refuge also in the past for these two species (Terhürne- Berson et al., 2004). Nevertheless, a significant portion of the flora currently present in this mountain range may be under threat for climate (and other) reasons (Thuiller et al., 2005; but see Scherrer & Körner, 2011).

Species distribution models are not free from the limitations and uncertainties widely discussed in the literature (Thuiller, 2004; Jiménez-Valverde et al., 2008; Thuiller et al., 2008). In addition to

management, some important factors involved in species' resistance to changes in distribution ranges are interactions between climate and resource availability (for instance, CO₂, see Keenan et al., 2011), phenotypic and genetic adaptation (Visser, 2008; Valladares et al., 2007), dispersal ability across fragmented landscapes (Opdam and Wascher, 2004), and biotic interactions, such as pathogens (Negrón et al., 2009). These factors may produce an under- or over- estimation of the future species suitability obtained from current climatic suitability. There is great uncertainty, however, about the ultimate effect of all these factors, and the resultant adaptation to new conditions is still subject to controversy (Davis & Shaw, 2001; Hamrick & Godt, 1996; Rehfeldt et al., 2002; Jump and Peñuelas, 2005; Kramer et al., 2010).

The extent of the distribution range considered in modeling may affect local estimations of extinction (Barbet-Massin et al., 2010; Thuiller et al., 2004). In our case, both species present a wide European distribution (Lahti & Lampinen, 1999) and the use of this expanded territory in the model would involve profiling a colder niche; more extinction would therefore be predicted in the Pyrenees. We argue that the use of our regional data is appropriate because of the importance of local adaptations (Hamrick & Godt, 1996; Sork et al., 2010), particularly in the case of beech (Kramer et al., 2010). Also, it could be argued that the use of constant variables over time (i.e. terrain) may lead to an overestimation of species persistence. It is common to explicitly ignore terrain variables in climate change in order to achieve a strict assessment of climatic effects on distribution (Araújo et al., 2006; Araújo and Luoto, 2007; Araújo and Guisan, 2006; Thuiller et al., 2006), but terrain variables play a key role in species distribution (Franklin, 1998; Leathwick, 1998; Hara, 2010; Austin & Van Niel, 2011). Other terrain variables not available for this study may significantly enhance the accuracy to our results. Soil cover, for example, is particularly important in mountain areas since mountain soils are often not sufficiently developed to sustain tree growth. The absence of this variable in the model may result in an overestimation of species persistence or an unrealistic upward altitudinal shift. For instance, Scherrer and Körner (2011) found that the terrain induced mosaic of environmental conditions buffers the flora of the Alps against climate change.

Co-occurrences do not track the whole environmental space of the species that compose a community, and community models are therefore a less general approach than the overlap of

individual species' responses. In general, a major argument for selecting the overlapping of individual models (OIM) is its greater capacity to model different situations (Guisan & Zimmermann, 2000; Baselga & Araújo, 2009). Although the model performance for a community may achieve good results, it seems less informative in terms of future niche differentiation than the OIM, which can assess future fitness at the level of an individual species. The analysis of niche differentiation, even across large regional scales, may provide information that enhances our understanding of local coexistence across environmental gradients (Silvertown, 2004).

On the other hand, profiling the realized niche of the community allows for a comparison between the environmental space of the observed co-occurrence and the environmental space of combinations (overlap) of the current individual species distributions. Similar concepts appear at the species level with equations such as the range/range potential ratio (Gaston, 2003). This has been used to evaluate species range filling on the European scale (Svenning and Skov, 2004). At the individual species level, it has been observed that current distribution of *Abies alba* and *Fagus sylvatica* is smaller than the potential realized environmental niche. Some authors state that low migration rates or even the existence of beech forests would have hampered the spread of *Abies alba* in this area during Holocene (Peñalba, 1994). However, despite the historical component, much of the differences between realized distribution and its potential environmental space may be attributed to biotic interactions, namely competition and facilitation processes.

According to our expectations, the realized niche of the community (modeled by the community model; CM) is spatially included in the pseudo-fundamental niche of the community (modeled by the overlap of individual species models; OIM) under present conditions. But this pattern differs in all the climate scenarios studied. Although both approaches predict similar declines in suitability, there is a relevant spatial disagreement. Although such differences are rooted in model parameterization, this highlights the relevance of the modeling approach (CM or OIM) in climate change analysis of the distribution of species co-occurrences. Further research should be undertaken using different modeling techniques and ensembles of forecasts in order to achieve more general results (Thuiller et al., 2009; Araújo and New, 2007).

We conclude that the climatic suitability of *Abies alba*, *Fagus sylvatica* and their mixed forests will led to a significant exposure to climate change, possibly leading to a decline, whichever modeling approach is chosen. Accuracy measures and future scenario trends are similar in both the community model and the overlap of individual models, but spatial projections differ, resulting in variations in the assessment of future climate-change effects on present distributions. We argue that the OIM is preferable for future scenarios, because of its broader generality, but the CM approach may provide information that determines the current shape in the environmental space occupied by co-occurrence, especially if compared to the full range of conditions in which the two species occur, either jointly or separately.

5.5 References

- Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, **1**, 95–111.
- Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.
- Barbet-Massin M, Thuiller W, Jiguet F (2010) How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? *Ecography*, **33**, 878–886.
- Ameztegui A, Coll L (2011) Tree dynamics and co-existence in the montane- sub-alpine ecotone: the role of different light-induced strategies. *Journal of Vegetation Science*, **22**, 1049–1061.
- Araújo MB, Thuiller W, Pearson RG (2006) Climate warming and the decline of amphibians and reptiles in Europe. *Journal of Biogeography*, **33**, 1712–1728.
- Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677–1688.
- Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, **16**, 743–753.
- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42–47.
- Austin MP, Van Niel KP (2011) Impact of landscape predictors on climate change modelling of species distributions: a case study with *Eucalyptus fastigata* in southern New South Wales Australia. *Journal of Biogeography*, **38**, 9–19.
- Bakkenes M, Alkemade JRM, Ihle F, Leemans R, Latour JB (2002) Assessing effects of forecasted climate change on the diversity and distribution of European higher plants for 2050. *Global Change Biology*, **8**, 390–407.
- Baselga A, Araújo MB (2009) Individualistic vs community modelling of species distributions under climate change. *Ecography*, **32**, 55–65.
- Beaumont LJ, Pitman AJ, Poulsen M, Hughes L (2007) Where will species go? Incorporating

new advances in climate modelling into projections of species distributions. *Global Change Biology*, **13**, 1368–1385.

- Beckage B, Osborne B, Gavin DG, Pucko C, Siccama T, Perkins T (2008) A rapid upward shift of a forest ecotone during 40 years of warming in the Green Mountains of Vermont. *Proceedings of the National Academy of Sciences of the U.S.A.*, **105**, 4197–4202.
- Benayas R, José M, Scheiner SM (2002) Plant diversity, biogeography and environment in Iberia: patterns and possible causal factors. *Journal of Vegetation Science*, **13**, 245–258.
- Benito Garzón M, Sánchez de Dios R, Sáinz Ollero H (2007) Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and Mid-Holocene. *Ecography*, **30**, 120–134.
- Benito Garzón M, Sánchez de Dios R, Sainz Ollero H (2008) Effects of climate change on the distribution of Iberian tree species. *Applied Vegetation Science*, **11**, 169–178.
- Botkin DB, Saxe H, Araújo MB et al. (2007) Forecasting the effects of global warming on biodiversity. *Bioscience*, **57**, 227–236.
- Brzeziecki B, Kienast F, Wildi O (1993) A simulated map of the potential natural forest vegetation of Switzerland. *Journal of Vegetation Science*, **4**, 499–508.
- Brzeziecki B, Kienast F, Wildi O (1995) Modelling potential impacts of climate change on the spatial distribution of zonal forest communities in Switzerland. *Journal of Vegetation Science*, **6**, 257–268.
- Chefaoui RM, Lobo JM (2008) Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modeling*, **210**, 478–486.
- Davis MB, Shaw RG (2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673–679.
- Elith J, Leathwick JR (2009) Species distribution models, ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, **40**, 677–697.
- Elith J, Graham CH, Anderson RP et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, **43**, 393–404.
- Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction

- errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.
- Franklin J (1998) Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*, **9**, 733–748.
 - Franklin J (2010) *Mapping Species Distributions: Spatial Inference and Prediction*. Cambridge University Press, Cambridge, UK.
 - Gaston KJ (2003) *The Structure and Dynamics of Geographic Ranges*. Oxford Univ. Press, Oxford, UK.
 - Gordon C, et al., (8 authors) 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, **16**, 147–168.
 - Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
 - Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
 - Hamrick JL, Godt MJW (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, **351**, 1291–1298.
 - Hara M (2010) Climatic and historical factors controlling horizontal and vertical distribution patterns of two sympatric beech species *Fagus crenata* Blume and *Fagus japonica* Maxim., in eastern Japan. *Flora*, **205**, 161–170.
 - Hirzel AH, Lay GL (2008) Habitat suitability modelling and niche theory. *Journal of Applied Ecology*, **45**, 1372–1381.
 - Holt RD (2009) Bringing the Hutchinsonian niche into the 21st century: ecological and evolutionary perspectives. *Proceedings of the National Academy of Science of the U.S.A.*, **106**, 19659–19665.
 - Hutchinson GE (1957) Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, **22**, 415–427.
 - IPCC (2007) Climate change 2007: climate change impacts, adaptation and vulnerability. In: Parry, M.L., Canziani, O.F., Palutikof, J.P., van der Linden, P.J., Hanson, C.E. (Eds.), *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental*

Panel on Climate Change 2007. Cambridge University Press, Cambridge, UK.

- Jiménez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, **14**, 885–890.
- Jump AS, Peñuelas J (2005) Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters*, **8**, 1010–1020.
- Jump AS, Hunt JM, Peñuelas J (2006) Rapid climate change-related growth decline at the southern range edge of *Fagus sylvatica*. *Global Change Biology*, **12**, 2163–2174.
- Jump AS, Mátyás C, Peñuelas J (2009) The altitude for latitude disparity in the range retractions of woody species. *Trends in Ecology and Evolution*, **24**, 694–701.
- Kearney M (2006) Habitat, environment and niche: what are we modelling? *Oikos*, **115**, 186–191.
- Keenan T, Serra-Diaz JM, Lloret F, Ninyerola M, Sabate S (2011) Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO₂ matters! *Global Change Biology*, **17**, 565–579.
- Kelly AE, Goulden ML (2008) Rapid shifts in plant distribution with recent climate change. *Proceedings of the Natural Academy of Sciences of the U.S.A.*, **105**, 11823–11826.
- Kramer K, Degen B, Buschbom J et al. (2010) Modelling exploration of the future of European beech (*Fagus sylvatica* L.) under climate change – range, abundance, genetic diversity and adaptive response. *Forest Ecology and Management*, **259**, 2213–2222.
- Lahti T, Lampinen R (1999) From dot maps to bitmaps – *Atlas Florae Europaeae* goes digital. *Acta Botanica Fennica*, **162**, 5–9.
- Leathwick JR (1998) Are New Zealand's *Nothofagus* species in equilibrium with their environment? *Journal of Vegetation Science*, **9**, 719–732.
- Lloret F, González-Mancebo JM (2011) Altitudinal distribution patterns of bryophytes in the Canary Islands and vulnerability to climate change. *Flora*, **206**, 769–781.
- Lloret F, Estevan H, Lobo P, Maisongrande P, Vayreda J, Terradas J (2007) Woody plant richness and NDVI response to drought event in Catalanian (NE Spain) forests. *Ecology*, **88**, 2270–2279.
- Lobo JM, Jiménez-Valverde A, Hortal J (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103–114.

- Macias M, Andreu L, Bosch O, Camarero J, Gutierrez E (2006) Increasing Aridity is enhancing silver fir (*Abies alba* Mill.) water stress in its south-western distribution limit. *Climatic Change*, **79**, 289–313.
- Meehl GA, Tebaldi C (2004) More intense, more frequent, and longer lasting heat waves in the 21st century. *Science*, **305**, 994–997.
- Miller J, Franklin J (2002) Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. *Ecological Modelling*, **157**, 227–247.
- Ministerio de Agricultura PyA (2007) Tercer Inventario Forestal Nacional. Ministerio de Agricultura PyA, Madrid, Spain.
- Negrón JF, McMillin JD, Anhold JA, Coulson D (2009) Bark beetle-caused mortality in a drought-affected ponderosa pine landscape in Arizona, USA. *Forest Ecology and Management*, **257**, 1353–1362.
- Ninyerola M, Pons X, Roure JM (2000) A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology*, **20**, 1823–1841.
- Nogués-Bravo D, Araújo MB, Errea MP, Martínez-Rica JP (2007) Exposure of global mountain systems to climate warming during the 21st century. *Global Environmental Change*, **17**, 420–428.
- Opdam P, Wascher D (2004) Climate change meets habitat fragmentation: linking landscape and biogeographical scale levels in research and conservation. *Biological Conservation*, **117**, 285–297.
- Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.
- Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, **133**, 225–245.
- Peguero-Pina JJ, Camarero JJ, Abadía A et al. (2007) Physiological performance of silver-fir (*Abies alba* Mill.) populations under contrasting climates near the south-western distribution limit of the species. *Flora*, **202**, 226–236.
- Peñalba MC (1994) The history of the Holocene vegetation in Northern Spain from pollen

analysis. *Journal of Ecology*, **82**, 815–832.

- Peñuelas J, Boada M (2003) A global change-induced biomes shift in the Montseny mountains (NE Spain). *Global Change Biology*, **9**, 131–140.
- Peñuelas J, Filella I, Comas P (2002) Changed plant and animal life cycles from 1952 to 2000 in the Mediterranean region. *Global Change Biology*, **8**, 531–544.
- Pope VD, Gallani ML, Rowntree PR, Stratton RA (2000) The impact of new physical parametrizations in the Hadley Centre climate model: HadCM3. *Climate Dynamics*, **16**, 123–146.
- Puddu A, Luisi N, Capretti P, Santini A (2003) Environmental factors related to damage by *Heterobasidion abietinum* in *Abies alba* forests in Southern Italy. *Forest Ecology and Management*, **180**, 37–44.
- Randin CF, Engler R, Normand S et al. (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, **15**, 1557–1569.
- R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, <http://www.R-project.org>.
- Rehfeldt GE, Tchebakova NM, Parfenova YI, Wykoff WR, Kuzmina NA, Milyutin LI (2002) Intraspecific responses to climate in *Pinus sylvestris*. *Global Change Biology*, **8**, 912–929.
- Schar C, Jendritzky G (2004) Climate change: hot news from summer 2003. *Nature*, **432**, 559–560.
- Scherrer D, Körner C (2011) Topographically controlled thermal–habitat differentiation buffers alpine plant diversity against climate warming. *Journal of Biogeography*, **38**, 406–416.
- Silvertown J (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution*, **19**, 605–611.
- Sork VL, Davis FW, Westfall R et al. (2010) Gene movement and genetic association with regional climate gradients in California valley oak (*Quercus lobata*) in the face of climate change. *Molecular Ecology*, **19**, 3806–3823.
- Svenning J, Skov F (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, **7**, 565–573.
- 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.

- Thomas CD, Cameron A, Green RE et al. (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thornthwaite CW (1948) An approach toward a rational classification of climate. *Geographical Review*, **38**, 55–94.
- Thuiller W (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020–2027.
- Thuiller W, Vayreda J, Pino J, Sabate S, Lavorel S, Gracia C (2003) Large-scale environmental correlates of forest tree distributions in Catalonia (NE Spain). *Global Ecology and Biogeography*, **12**, 313–325.
- Thuiller W, Brotons L, Araújo MB, Lavorel S (2004) Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, **27**, 165–172.
- Thuiller W, Lavorel S, Araújo MB, Sykes MT, Prentice IC (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the U.S.A.*, **102**, 8245–8250.
- Thuiller W, Lavorel S, Sykes MT, Araújo MB (2006) Using niche-based modeling to assess the impact of climate change on tree functional diversity in Europe. *Diversity and Distribution*, **12**, 49–60.
- Thuiller W, Albert C, Araújo MB et al. (2008) Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 137–152.
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD. A platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369–373.
- Valladares F, Gianoli E, Gómez JM (2007) Ecological limits to plant phenotypic plasticity. *New Phytologist*, **176**, 749–763.
- Visser ME (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society of London B: Biological Sciences*, **275**, 649–659.
- Walther GR, Post E, Convey P et al. (2002) Ecological responses to recent climate change.

Nature, **416**, 389–395.

- Williams JW, Jackson ST (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, **5**, 475–482.

6. ON FACING POTENTIAL NICHE CHANGES

STUDY: ‘Knowing the origin to predict the end: spatial projections in *Rhaponticum repens* (Compositae) suggests considering its invasion provenance.’

6.1 Introduction

Invasive species are amongst the most relevant world threats to ecosystems and biodiversity (Ecosystem Millennium Assessment, 2005) and with an increasing globalized world where vectors of invasion multiply; a good understanding of the environmental conditions prone to invasion across different biological realms is of paramount interest.

Central to the concept of environmental suitability of a species is niche theory: a set of conditions in which a species can develop and reproduce. First coined by Grinnell (1917) and later conjectured by Elton (1930) to food-webs, the operational framework of the theory is attributed to Hutchinson (1957) which defined the ecological niche as the n-multidimensional space where species can maintain its living. Further conceptual developments have been explored by (Soberón, 2007) which linked such multidimensional spaces with mechanisms using the BAM diagram, which interprets species distribution according to three interrelated dimensions: biotic (B), abiotic (A), geographic (M).

Environmental suitability of a species can be measured via environmental niche models (ENM, also termed species distribution models). This correlative modeling approach assesses the relationship between species occurrence or abundance to environmental variables thought to influence species fitness and distribution (Franklin 2010). The statistical method used to construct the model varies from regression fitting to machine learning approaches, which influence model outcomes and interpretation (Elith & Graham, 2009; Jiménez-Valverde et al., 2008) , together with other factors and assumptions that make environmental niche models more than a simple application of an algorithm (Elith et al., 2010) . Overall, these models only constitute a preventive

screening phase of invasion risk assessment (Drake & Lodge, 2006; Peterson, 2003), but they gained popularity given that the costs of eradication and potential economic consequences (Pimentel et al., 2005).

In the case of invasive species, environmental niche models have proved to be very useful to determine new areas of invasions (Roura-Pascual et al., 2004) and overall helped on assessing invasive risk assessment and understand environmental drivers of invasion (Ficetola et al., 2007; López-Darias & Lobo, 2008; Capinha & Anastácio, 2011). Traditional applications consist in estimating the environmental niche in native ranges and subsequently project them to new invasive areas (but see Fitzpatrick et al., (2007)), which has lead to outstanding results in some cases (Richardson & Thuiller, 2007; Ibáñez et al., 2009). The rationale of such procedure relies on the assumptions of niche stasis or slow niche evolution in ecological time, which has its homologue evolutionary concept in the niche conservatism (Wiens et al., 2010). In other words, we assume that the niche does not change and we are able to capture it via native range distribution, which is more likely to be in equilibrium with climate than in the invasive range (Peterson, 2005).

Unfortunately but interestingly the general picture is by far more complex, challenging a straightforward application of the native niche estimation projected to potential invasive areas. Niche shifts have been reported to occur during invasion (Broennimann et al., 2007; Medley, 2010; Urban et al., 2007; da Mata et al., 2010) and several processes may be playing an important role. In some cases it is erroneous to assume species' niche stasis (at least the realized niche) or slow evolution, which could be the case of genome reduction during invasion (Lavergne et al., 2010), genetic bottleneck (Golani et al., 2007), mutations (Phillips et al., 2008) or hybridization (Hall et al., 2006; Schierenbeck et al., 2009; Mukherjee et al., 2011). In addition, several decisions undertaken during the modeling process (e.g., equilibrium, algorithm used, etc.) may also underestimate or overestimate niche shifts and potential invasion area (Jiménez-Valverde et al., 2011).

Russian knapweed (also known as creeping Knapweed) [*Rhaponticum repens* (L.) Hidalgo] is a rhizomatous perennial plant native from Turkey throughout Central Asia to China (Koloren et al., 2008) that has become a noxious weed in several countries as USA, Canada, Argentina, Germany, Italy, Western Australia and recently has been reported from eastern Iberian Peninsula (López

Alvarado et al., 2011; see Figure 6.1). Russian knapweed was initially introduced to North America in the early 1900's, primarily as a contaminant of Turkestan alfalfa (*Medicago sativa*) seed (Watson, 1980) and possibly sugarbeet (*Beta vulgaris*) seed (Maddox et al., 1985). It is a serious weed pest which can establish and persist in a variety of environments, such as disturbed grassland, croplands, irrigation ditches, roadsides, shrubland communities and riparian woodlands (Carpenter & Murray, 1999). The main method of spread for Russian knapweed (which produces relatively few seeds and lacks effective mechanisms for seed dispersal) is by adventitious buds on a creeping perennial root system (Watson, 1980). This species contains an allelopathic polyacetylene compound which inhibits the growth of competing plants (Watson, 1980; Stevens, 1986). This allelopathic effect, combined with effective clonal growth, allows *Rhaponticum repens* to quickly colonize and dominate new sites. Preliminary molecular analysis of *Rhaponticum repens* from diverse geographical origins has evidenced that specimens from Iberian Peninsula could be related to USA populations (Nearctic realm) rather than with those of central Asia (native range), indicating that invasion in the Iberian Peninsula can be the result of a secondary invasion (López Alvarado et al., 2011).

Under the urgent need to determine invasion risk to undertake preventive actions (preventive assessment), it is crucial to consider and anticipate potential scenarios of niche shifts during invasion. In the present study we assess the prediction capacity and environmental niche similarity of ENMs in spatial projections of *Rhaponticum repens* into an invaded area (Iberian Peninsula). We hypothesize that knowing the origin of invasions should translate into better predictions of invasion due to lower probability of niche shift, at least in the initial phase. We interpret these results in light of the decisions taken during the modeling process by using several ENMs calibrated in both different biogeographical realms and the world (using all information available) (Broennimann & Guisan, 2008).

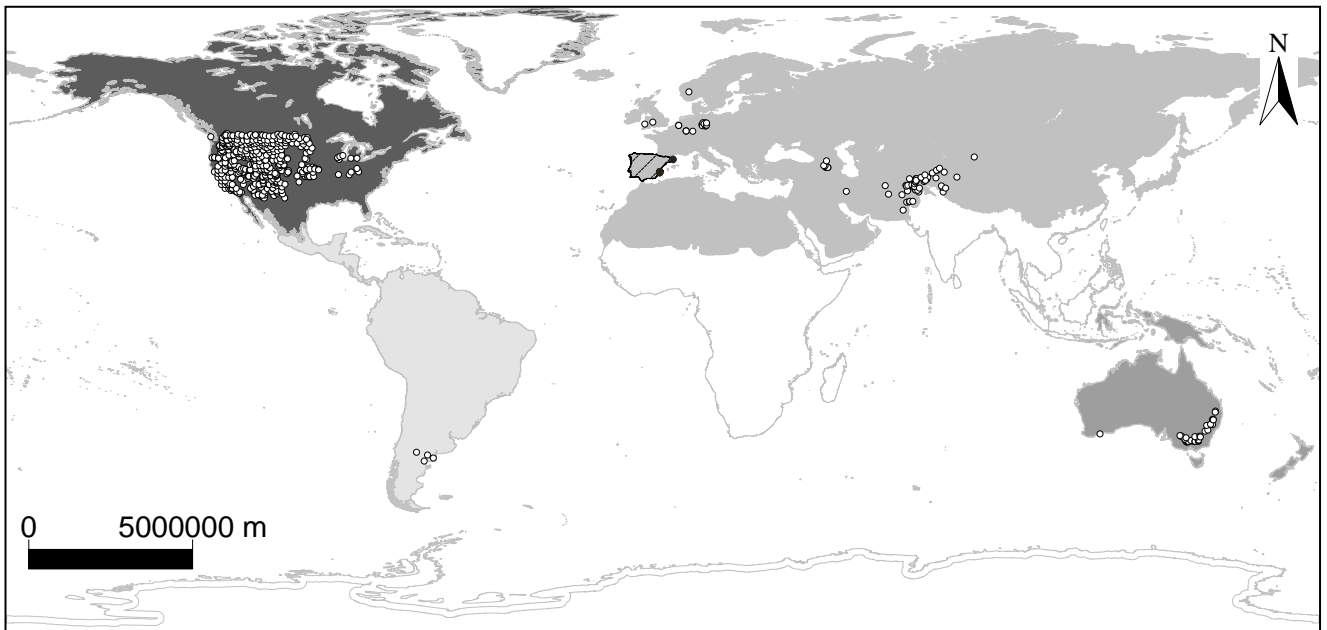


Figure 6.1. Distribution of *Rhaponticum repens* (L.) in different realms of the world (different shaded areas). Black dots indicate main invaded locations (N=4) in the Iberian Peninsula (dashed lines).

6.2 Material and methods

Environmental niche models (ENM) have been constructed in their native and other invaded distributional areas of the world, encompassing 4 biological realms (World Wildlife Fund; (Olson et al., 2001): Palearctic ($N_{\text{native}}=64$; $N_{\text{invasive}}=25$), Nearctic ($N=497$), Neotropical ($N=4$), Australian ($N=36$). An additional global model has been calibrated using all records available (World model; Broennimann & Guisan, 2008) (see model data in Figure 6.2). Data of species presence was extracted from the global biodiversity information facility (GBIF) database and subsequently data filtering using quality control processes (geopositional congruence) to avoid possible planimetric errors. Because of the relatively recent systematic studies, *Centaurea repens* has been transferred to the genus *Rhaponticum* as *Rhaponticum repens* (Hidalgo et al. 2006), which could lead to the underestimation of records, we ensured that previous species names (*Acroptilon picris*, *Centaurea repens*, *Acroptilon subdentatum*, *Acroptilon repens*, *Acroptilon obtusifolium*, *Carduus picris*,

Acroptilon angustifolium, , *Serratula picris*, *Acroptilon serratum*, *Centaurea picris*) where explicitly included. In the case of Neotropical model, which account for very low number of presences, model results have been built and projected but interpreted with caution, although species distribution models have shown good performance under such circumstances (Pearson et al., 2007), at least with sophisticated algorithms (Hernandez et al., 2006). All other models account for acceptable number of occurrences ($N \geq 25$, (Araújo et al., 2005)).

Climatic data was extracted from the Worldclim database (Hijmans et al., 2005) at 2.5 arc minutes resolution (approximately 5 km at the Equator). In order to avoid excessive over-parametrization of models possibly leading to overfitting (Warren & Seifert, 2011), we selected 6 bioclimatic variables that have strong influence over plant physiology and distribution: Annual precipitation (Pann) , mean annual temperature (Tmean), maximum temperature of the warmest month (Tmaxwarm), minimum temperature of the coldest month (Tmincold), precipitation of the wettest month (Pwet) and precipitation of the driest month (Pdry).

Two modeling techniques have been employed: a profiling technique (presence-only; Bioclim (Busby, 1991) and context dependent technique (presence – background; Maxent (Phillips et al., 2006). Bioclim algorithm profiles current climatic situations where the species is present without any consideration of the species absence (see Figure 6.2). It calculates the mean and standard deviation of species occurrences for each environmental variable and characterizes a given environment as suitable, marginal or unsuitable depending on the number of variables within a standard deviation criterion. We used the implementation in OpenModeler software (de Souza Muñoz et al., 2011), which uses a default standard deviation cutoff of 0.674.

Maxent characterizes both the occurrences of the species and the landscape of occurrence (background, see Figure 6.2) and minimizes the entropy between the two probability densities (occurrence and background). We used Maxent v.3.3.3k using a conventional approach of allowing automatic features selection, with a regularization parameter of 1. We subsampled occurrences setting 30% of occurrences for validation and using 4 replicates. Because Maxent also characterizes the background, such algorithm implies that the delimitation of the available environment affect its results and interpretation (Barve et al., 2011). Therefore, we further developed two background

environment scenarios (see Figure 6.2): (1) considering the biological realm as potential accessible or accessed area (background), and (2) restricted accessible area to potential dispersal observed in each invasion. In the latter case, available habitat is constrained by a distance kernel derived from the cumulative distance (D_i) between current invasive locations (Václavík & Meentemeyer, 2009). Such measure is suitable to incorporate dispersal limitations in invasive ENM without detailed consideration of dispersal traits (Allouche et al., 2008).

We projected every model calibrated in different regions to the biogeographical space to a recent invaded area (Iberian Peninsula) obtaining a map of environmental suitability of *Rhaponticum repens* that ranges from 0 (unsuitable) to 1 (suitable). We further tested the approach of using all available data by calibrating another model for the entire world (Broennimann & Guisan, 2008). Therefore, a total of 18 environmental suitability maps were evaluated representing different combinations of modeling algorithm, invasion provenance and background scenarios.

In order to distinguish suitable and unsuitable invasion conditions a threshold has to be set and there exist many procedures to calculate it (Freeman & Moisen, 2008; Liu et al., 2005). Despite being a great source of uncertainty in modeling projections (Nenzén & Araújo, 2011), it is of a general agreement that the choice of the threshold relies mainly on the goal of the study (Jiménez-Valverde & Lobo, 2007). In general, species distribution models tend to use a procedure where thresholds optimize presence and absence/background predictions. However, in invasive risk assessment it has been argued that the choice of a threshold should be based on the maximization of the true positive rate (TPR, presences correctly predicted) because the costs of extraction surpass the costs of prevention (Jiménez-Valverde et al., 2011). Therefore, we investigated the effect of two threshold criteria: equate entropy of original and thresholded distributions (EET; optimized threshold where entropy in both distribution is kept equal) and minimum training presence (MTP; conservative threshold where all presences are predicted). Differences in predictions have been interpreted through the percentage of correct predictions (TPR) in the new invasions, the environmental suitability index of the new invasions and the total area climatically suitable for invasion.

We further assessed projections in the Iberian Peninsula of different models by measuring niche overlap with three different metrics: (Schoener, 1968) overlap index (D), a modified Hellinger distance index (I) (Van der Vaart, 2000) and a relative rank index (RR) (Warren & Seifert 2011). Such measures provide a quantitative approach to measure spatial congruence and similarity of projections (see Warren et al., (2008) for further details). D-index quantifies the degree of similarity between two probability distribution over a geographical space ranging from 0 (no overlap) to 1 (identical projection) using the differences in probabilities in the same locations. Nevertheless, it is unwarranted to assume that such probabilities are proportional to species density or a measure of relative use (Warren et al., 2008) and therefore authors proposed in addition the use of a 0 to 1 transformation of Hellinger's distance (I), which merely compares probability of distributions. Additionally, we used a third metric of relative rank (RR) proposed by Warren & Seifert (2011), in which probabilities in cells are compared according to their ranking in their projections. Using a similar example of the authors, comparing two projections in which differences in probabilities are proportional across the entire geographical space, would yield low similarity in their I and D similarities but RR would indicate a complete correspondence. Thus, such measure indicates given the relative geographical ranking despite raw differences in probabilities for each location. In our study, this metric is of special interest because it draws the qualitative spatial projections of invasion risk. The entire methodological sequence is illustrated in Figure 6.2.

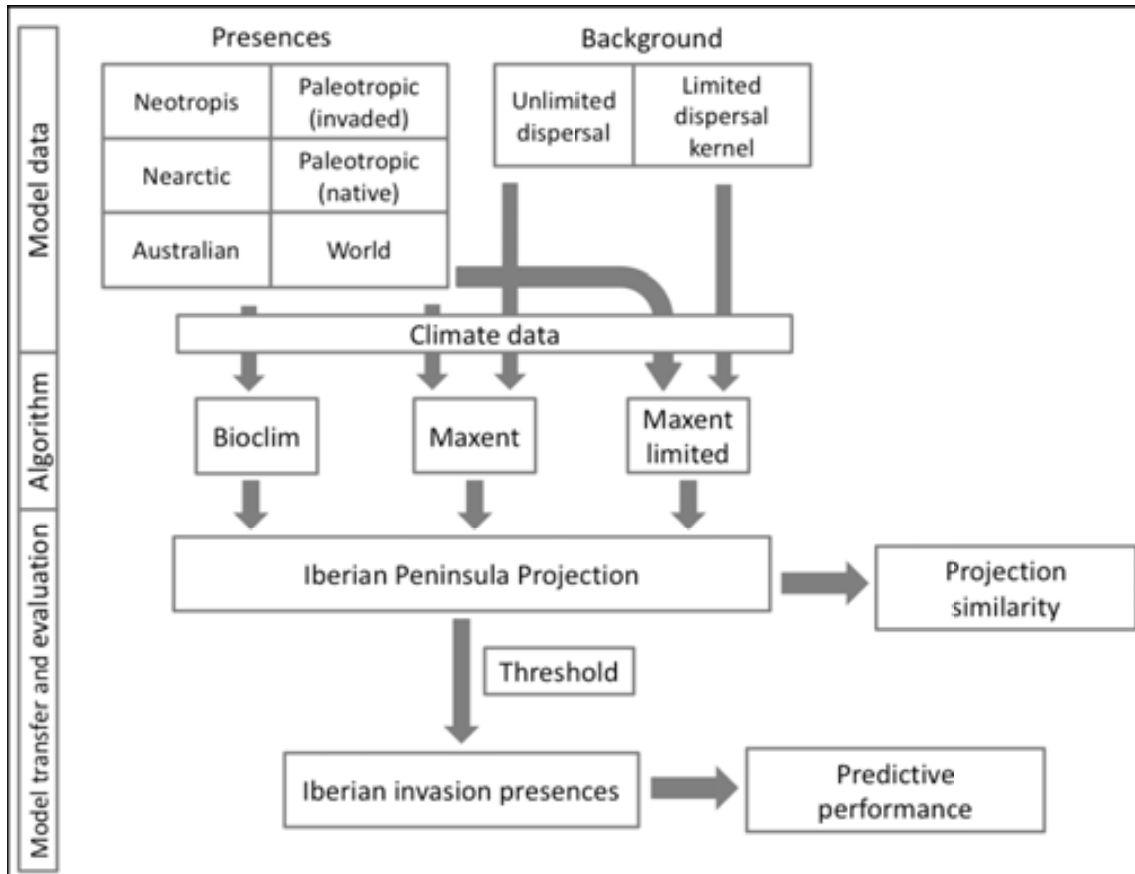


Figure 6.2. Sequence of methods used.

6.3 Results

The genetically identified area of provenance (Nearctic, NA) modeled with Bioclim and Maxent without dispersal constraints, showed the best stable invasive predictive power (Figure 6.3): all new invasions were predicted regardless of the chosen threshold. Nevertheless, modeling the complete range of the species in the world or in the Australian realm (with MTP threshold) achieved nearly as good results as using the origin of invasion for calibrating models (Nearctic). In our case, the traditional approach of projecting the native range to the invasive range did not yield good results with a mean percentage of correct predictions of 50%, but no combination of parameters achieved the desired predictive power of 100% (Figure 6.3).

All evaluated models showed acceptable performance results in their respective area of calibration with AUC values ranging from 0.785 to 0.986, with the exception of the Neotropical dispersal constrained model (0.512, see Table 1). We argue that in that case the relatively low number of presences clustered in a constrained region does not make climate a good predictor of presence nor a reliable model. Concurring with other studies (Hernandez et al., 2006, Wisz et al., 2008) , the limitation of the available area (dispersal distance constraint) and low number of presences lowered accuracy values (AUC), although they still qualify for acceptable models (using Swets (1988) classification).

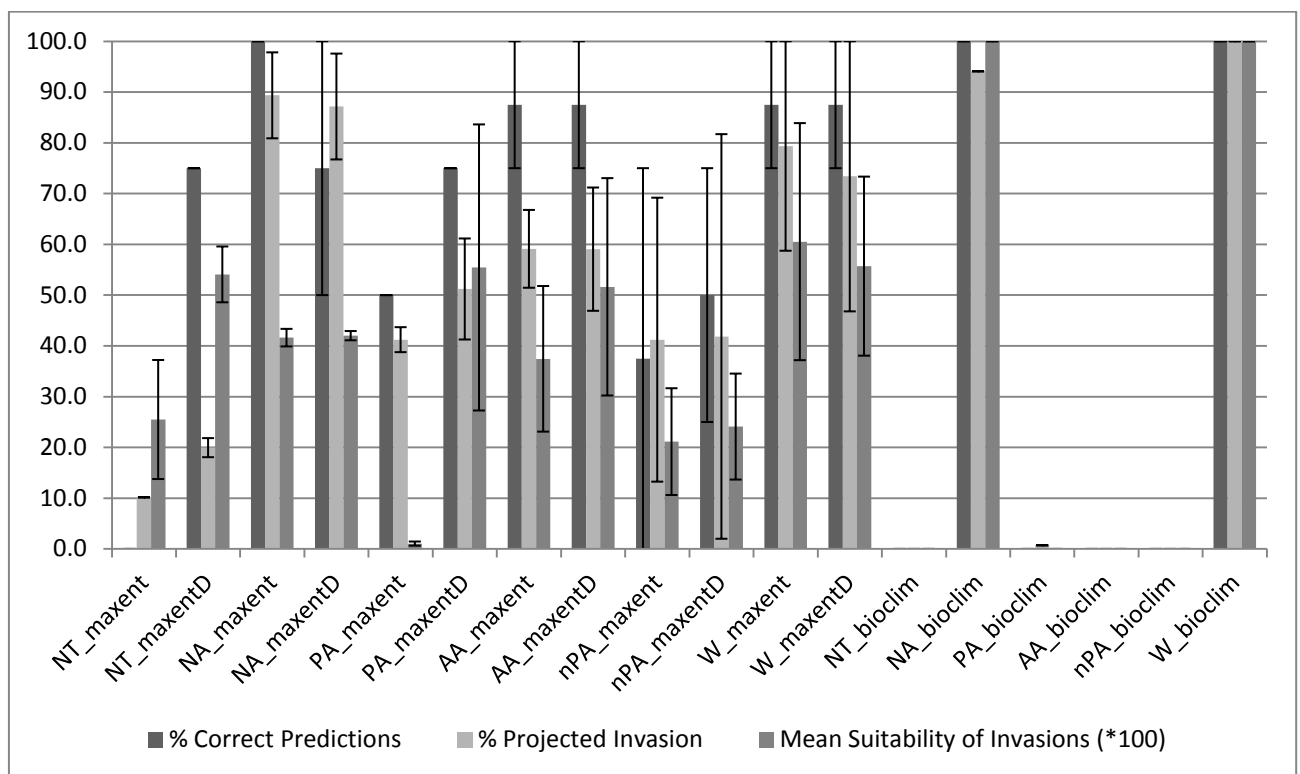


Figure 6.3. Results of projections to the Iberian Peninsula, according to different provenance models. Mean suitability of invasions indicates the average of the probability values in invaded locations. Percentage of projected invasion refers to the area predicted to be climatically suitable for invasion. Percentage of correct predictions indicates the number of current invasions predicted in each projection. Error bars indicate differences between equal entropy threshold (upper boundary) and minimum training presence threshold (lower boundary). Bars indicate average between the two threshold cases.

MODEL	AUC
Neotropical	0.98
Neotropical D	0.512
Nearctic	0.899
Nearctic D	0.843
Paleartic (invasive)	0.986
Paleartic (invasive) D	0.844
Australian	0.961
Australian D	0.899
Paleartic (Native)	0.898
Paleartic (Native) D	0.785
World	0.931
World D	0.811

Table 6.1. Provenance Maxent models' accuracy, using the area under the receiving operating curve (ROC, (Fielding and Bell 1997)). D indicates whether background has been constrained using maximum dispersal kernel (Allouche et al. 2008).

Bioclim algorithm projected two contrasting scenarios: huge climate-based invasion risk (Nearctic and World models) or no risk at all (Palearctic, Neotropic, native range and Australian model) (Figure 6.3). These two outputs are influenced by the low flexibility of the algorithm, which produces a hard classification (no gradient or soft boundaries) of projections that tend to predict 'all or none' areas of invasion. In the case of the world provenance model, although it fully predicts all invasions (Figure 6.3), the model is of limited use because the entire Iberian Peninsula is predicted to be at the same risk, therefore the model is not capable to prioritize any area over another (Figure 6.4).

Maxent algorithm produces much variable outputs and the threshold set largely influence the success on predicting new invasions (Figure 6.3). As expected, optimized error threshold (EET) produces poorer prediction rates than conservative (MTP) threshold. For instance, using MTP (conservative) threshold criteria, Australian model predicts all invasions in the Iberian Peninsula and assign high probability values to invasions although it predicts less suitable areas than the Nearctic model (invasion provenance). In the case of the World model using the same threshold criteria, all invasions are also correctly predicted but they result uninformative since the entire Iberian Peninsula is predicted to be at risk (Figure 6.4). Interestingly in the Nearctic model (invasion provenance) without dispersal constraints, invasion predictive power is not affected by the threshold criteria used. Using constrained dispersal background in models did not show a straightforward effect on predictive power, nor the amount of area predicted nor probability values (Figure 6.3). For instance, in the Nearctic model (invasion provenance) yielded poorer invasive predictive power when applying such distance constraint whereas it increased in the case of the native range.

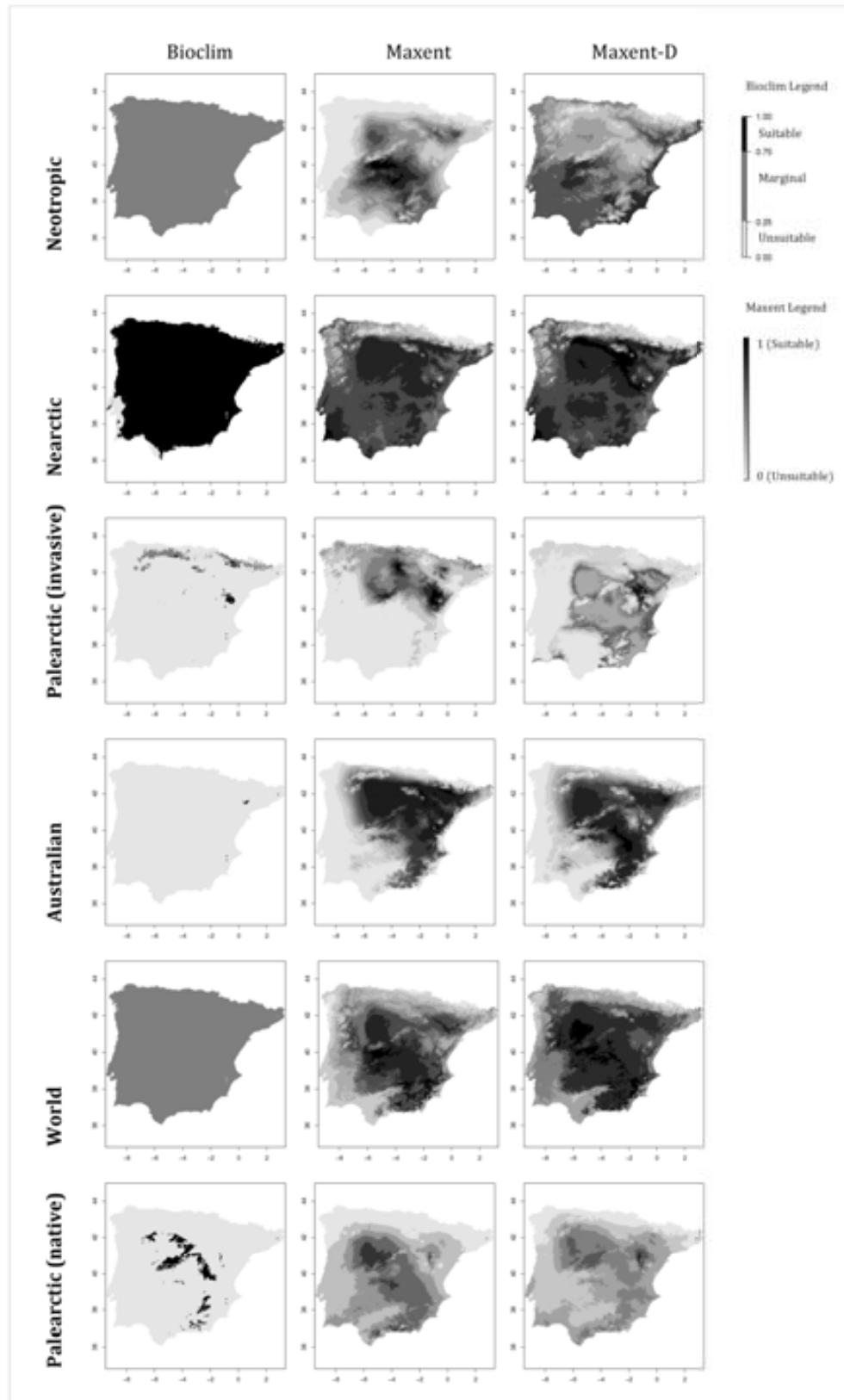


Figure 6.4. Spatial projections of invasions of *Rhaponticum repens* in the Iberian Peninsula, according to different invasion provenance

In general, spatial projections to the Iberian Peninsula showed markedly differences between modeling algorithms (Figure 6.4). As expected, Maxent generally provides complex patterns whereas most of Bioclim models provide simple classes due to the differences in the algorithm nature. Most projections indicate a relatively higher risk in the North-East and East of the Iberian Peninsula (Figure 6.4), but this is by far not a general obvious pattern. In general, projections agree on defining high mountain regions as unsuitable, but the role in coastal and inland environments is highly depending on models. Quantitative measures of niche overlap did not show a straightforward pattern although the majority of tests showed weak to null similarity among projections (Table 6.2). In line with results of invasion predictive power, niche overlap between native (nPA) and provenance (NA) models only reaches 0.53 in relative rank, although relatively similar for D and I index (Table 6.2). Such results imply that both projections predict a similar degree of regional invasion climatic suitability, but ranked differently in space. Interestingly, spatial mismatch is aggravated in the case of the world model (W), where niche overlap does not reach 0.50 in RR compared to projections of any other provenance model.

As expected, similarity measured by D and I show little differences because they rely on similar quantitative measures of probability of presence (Warren et al. 2008). In all cases, they present systematically higher values than RR. Such results point out that even if the projections may partially agree in their scoring of suitability, they present different ranking of areas suitable for invasion (e.g. prioritization of invasive areas). Adding dispersal constraints did not affect niche overlap tests univocally: in some instances overlap was high (Australian and Nearctic model) and in other models caused spatial mismatch in projections (Neotropic model).

Table 6.2. Niche overlap of provenance models in the Iberian Peninsula measured by relative rank (RR), Schoener index (D) and modified Hellinger's distance (I) (see Warren et al. 2008; Warren and Steifert 2011 for further details). Model codes: NT (Neotropical), NA (Nearctic), PA (Palearctic), AA (Australian), nPA (nativePalearctic), W (world). Suffix ' _D' indicate background constrained by maximum dispersal kernel (Allouche et al. 2008). Bold numbers indicate clear similarity (all RR>0.6 and D>0.6 and I>0.6) and shaded cells indicate spatial i congruence (RR<0.5, D>0.5, I>0.5).

	NT			NT_D			NA			NA_D			PA			PA_D			AA			AA_D			nPA			nPA_D			W			W_D		
	RR	D	I	RR	D	I	RR	D	I	RR	D	I	RR	D	I	RR	D	I	RR	D	I	RR	D	I	RR	D	I	RR	D	I	RR	D	I			
NT				0.47	0.65	0.84	0.51	0.68	0.87	0.50	0.68	0.87	0.31	0.35	0.52	0.62	0.68	0.86	0.55	0.63	0.83	0.63	0.70	0.88	0.62	0.74	0.89	0.51	0.67	0.87	0.37	0.73	0.84	0.39	0.79	0.96
NT_D							0.39	0.83	0.97	0.39	0.83	0.97	0.17	0.33	0.55	0.39	0.55	0.75	0.25	0.51	0.75	0.32	0.56	0.80	0.43	0.69	0.88	0.33	0.69	0.90	0.45	0.84	0.96	0.57	0.89	0.98
NA										0.83	0.97	1.00	0.42	0.42	0.62	0.45	0.58	0.78	0.53	0.64	0.83	0.53	0.68	0.87	0.55	0.75	0.91	0.61	0.81	0.95	0.47	0.86	0.97	0.56	0.88	0.98
NA_D													0.43	0.42	0.62	0.45	0.57	0.77	0.52	0.64	0.83	0.52	0.68	0.87	0.53	0.74	0.91	0.58	0.80	0.95	0.46	0.86	0.97	0.35	0.43	0.64
PA																0.47	0.47	0.64	0.57	0.61	0.78	0.47	0.52	0.71	0.31	0.33	0.48	0.44	0.45	0.62	0.29	0.56	0.72	0.51	0.63	0.80
PA_D																			0.59	0.65	0.83	0.61	0.68	0.84	0.51	0.61	0.78	0.53	0.63	0.80	0.29	0.67	0.78	0.57	0.68	0.87
AA																						0.81	0.88	0.98	0.49	0.61	0.78	0.57	0.72	0.86	0.31	0.67	0.82	0.63	0.73	0.90
AA_D																									0.56	0.69	0.84	0.62	0.76	0.90	0.33	0.71	0.85	0.65	0.79	0.92
nPA																												0.66	0.81	0.95	0.48	0.78	0.92	0.60	0.81	0.95
nPA_D																															0.43	0.79	0.93	0.44	0.85	0.97
W																																				
W_D																																		0.63	0.73	0.88

6.4 Discussion

Our results support the hypothesis that calibration of *Rhaponticum repens* in their invasive origin (Nearctic) translates into better and less-parameter dependent projections of invasions in the new invasive range (Iberian Peninsula) than native range models. However, we acknowledge these results may not be applicable to all invasive situations: different phases of invasion may be taking place in different geographical regions, hence affecting climatic equilibrium status and hampering a good climatic profiling of the species. In fact, ENMs assume equilibrium with climate, which is not always the case in neither invasions (Roura-Pascual et al., 2009) nor species distributions (Araújo & Pearson, 2005). Václavík & Meentmeyer (2009) demonstrated that calibrating ENMs at early stages of invasion clearly underpredicts the extent of potential invasive environments, but such effect decreases at intermediate to large climatic equilibrium situations. In *Rhaponticum repens*, most introductions took place more than 90 years ago (date records extracted from GBIF database), except for the Neotropic realm where databases date introductions in 1972. Therefore, we considered that ENMs in this case are not biased due to their low climatic equilibrium status, although such factor could be significant in the Neotropic realm model.

In addition to equilibrium assumptions, model calibration in a geographic area different than the projection poses several challenges to the use of ENMs and termed as model transferability in space (Elith & Leathwick, 2009). In the developed models, new environmental conditions were not found and therefore we did not encounter climatic extrapolation issues (see Appendix 3.1). However, we did not check for biotic interactions that could potentially reduce or expand the climatic range in model calibration (Godsoe & Harmon, 2012), hence reducing transferability of ENMs. Also, the modeling technique used influence the degree of transferability to other regions. We used two contrasted modeling approaches to determine invasions: envelope modeling (presence only, Bioclim) and machine learning (presence-background, Maxent). The two of them provided different but complementary information. Bioclim algorithm showed that invasions in the Iberian Peninsula were climatically near the situation in the Nearctic presences. Derived maps from such models however are very simple and easily overestimate or underestimate suitable area. On the contrary, Maxent models provided informative spatialization and relative similarities between

Iberian and other invasions. Heikkinen et al. (2011) showed that Maxent achieved good performance in transferability among machine-learning techniques, although the authors also emphasize that in general transferability for plant ENMs is generally lower than for other species. Altogether, our results emphasize that ensemble modeling is highly recommended to reduce technique uncertainty (Araújo and New, 2007); but see also some warnings (Elith et al., 2010).

Besides modeling technique, the invasion predictive power showed significant interaction with other modeling decisions. Applying different thresholds showed that the most stable was the Nearctic provenance model, but it is important to stress that if a relatively lower threshold criterion is chosen other provenances could have achieved the same predictive power. Similarly and especially meaningful is the role of background or pseudo-absence selection. The statistically determined dispersal area, used as background, lowered the accuracy of the models of *Rhaponticum repens* in their calibration region and did not improve models in the invasive region. Indeed is the selection of background/pseudo-absence region one of the major challenges in species distribution models (Anderson & Raza, 2010; Chefaoui & Lobo, 2008). Barve et al. (2011) identified three approaches to identify background regions: biotic regions, niche reconstructions (hindcasting to measure potential distributional areas) and full dynamic approaches (estimations of dispersal potential through simulations). In this study, the buffer area selection used may picture a phase of invasion that does not suite the assumptions of pseudo-equilibrium of ENMs. It is likely that invasive situations require better simulated dispersal in order to characterize the potential area of distributions.

Other factors that we did not explore are the variables used and number of occurrences, which number and kind may also have a strong influence in models. Warren & Seifert (2011) showed that intermediate levels in the number of variables were preferred to avoid overfitting in Maxent. Given the global scale of our analysis and the prospective nature of the projections we refused the use of land use variables. Although very relevant, using only-climate variables allow for the determination of a first abiotic context where invasion could take place, whereas land use variable could be explored afterwards as a risk multiplier effect or at a more local scale (Ibañez et al., 2009; Santos et al., 2011). In this analysis, the relatively different number of presences may largely affect our results. In fact, we acknowledge that we cannot identify true niche differences unless our model

building does not account for these differences using, for instance, resampling methods. However, we rather used this study as a potential application of world versus provenance realm. Future studies being undertaken take into consideration both such sampling issues.

Above all the mentioned assumptions and methodological challenges, one of the biggest transferability issues is that species niches may change. Broennimann et al. (2007) demonstrated that a shift of the observed climatic niche of the invasive Spotted Knapweed (*Centaurea maculosa* Lam.) occurred between native and non-native ranges. However in a large-scale analysis, (Petitpierre et al., 2012) reported that niche shifts are rather the exception than the rule for terrestrial plant invaders. Often, attributed niche shifts may be explained by niche unfilling in the native range (via biotic or dispersal constraints) or because such environments are not found in native ranges. Therefore, the authors evidenced that identification of niche shifts poses several conceptual and methodological challenges (but see Broennimann et al. (2007) and Warren et al. (2008)).

It would be misleading to interpret our results as niche shifts in *Rhaponticum repens* invasions, but our add-hoc analysis pointed out that, given the possibility of climatic niche shifts, considering models' projections from several provenances may develop several scenarios in preventive risk assessment of invasions. Such projections evidenced that a large area of the Iberian Peninsula is climatically suitable for *Rhaponticum repens*. Models' projections mostly differ into what ranking is given between the coastal and the inland regions, being the best predictive projections those ranking higher for coastal conditions. We further hypothesize that the combination of climatic niche of invasive species provenance and introduction effort should increase our predictive power and understanding of invasions (Herborg et al., 2007), by developing invasive hybrid models (Hall et al., 2006). In a broader context and also in line with our results, using all available records with a low threshold criterion may prove to be to predict new invasions, fully concurring with recommendations reviewed by Jimenez-Valverde et al. (2012) and Broennimann & Guisan (2008). However, its effectiveness may rely more on decisions of the modeler. We recommend the use of alternative spatial scenarios of provenance to assess potential introductions via niche shift. However, we are still on the quest to find an ensemble of such projections that may facilitate depicting invasion risk and potential niche change at the same time.

6.5 References

- Allouche O, Steinitz O, Rotem D, Rosenfeld A, Kadmon R (2008) Incorporating distance constraints into species distribution models. *Journal of Applied Ecology*, **45**, 599-609
- Anderson RP, 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 MB, Pearson RG, Thuiller W, Erhard M (2005) Validation of species–climate impact models under climate change. *Global Change Biology*, **11**, 1504-1513
- Araújo MB, Pearson RG (2005) Equilibrium of species distributions with climate. *Ecography*, **28**, 693-695
- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, **22**, 42-47
- Barve N, Barve V, Jiménez-Valverde A, et al. (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modeling*, **222**, 1810-1819.
- Broennimann O, Treier UA, Müller-Schärer H, Thuiller W, Peterson A, Guisan A (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701-709
- Broennimann O, Guisan A (2008) Predicting current and future biological invasions: both native and invaded ranges matter. *Biology Letters*, **8**, 585-589
- Busby J (1991) BIOCLIM-a bioclimate analysis and prediction system. *Plant Protection Quarterly*, **6**, 8-9
- Capinha C, Anastácio P (2011) Assessing the environmental requirements of invaders using ensembles of distribution models. *Diversity and Distributions*, **17**, 13-24
- Carpenter, Alan T.; Murray, Thomas A. 1999. Element Stewardship Abstract for *Acroptilon repens* (L.) De Candolle/(*Centaurea repens* (L.)): Russian knapweed. In: Weeds on the web: The Nature Conservancy Wildland Invasive Species Program, [Online]. Available: <http://tncweeds.ucdavis.edu/esadocs/acrorepe.html> [2002, January 11].

- Chefaoui RM, Lobo JM (2008) Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling*, **210**, 478-486
- Drake JM, Lodge DM (2006) Forecasting potential distributions of nonindigenous species with a genetic algorithm. *Fisheries*, **31**, 9-16
- Elith J, Graham CH (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, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677-697
- Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods in ecology and evolution*, **1**, 330-342
- Elton, CS (1930) *Animal Ecology and Evolution*. Clarendon Press, Oxford.
- Ficetola GF, Thuiller W, Miaud C (2007) Prediction and validation of the potential global distribution of a problematic alien invasive species—the American bullfrog. *Diversity and Distributions*, **13**, 476-485
- Fielding AH, Bell JF (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.
- Freeman EA, Moisen GG (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modeling*, **217**, 48-58
- Fitzpatrick MC, Weltzin JF, Sanders NJ, Dunn RR (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
- Godsoe W, Harmon LJ (2012) How do species interactions affect species distribution models?. *Ecoraphy*, - doi: 10.1111/j.1600-0587.2011.07103.x
- Golani D, Azzurro E, Corsini-Foka M, Falautano M, Andaloro F, Bernardi G (2007) Genetic bottlenecks and successful biological invasions: the case of a recent Lessepsian migrant. *Biology Letters*, **3**, 541-545

- Grinnell J (1917) Field tests of theories concerning distributional controls. *American Naturalist*, **51**, 115-128. doi: 10.1086/279591
- Hall RJ, Hastings A, Ayres DR (2006) Explaining the explosion: modelling hybrid invasions. *Proceedings Royal Society Biology: Biological Sciences*, **273**, 1385-1389
- Herborg LM, Jerde CL, Lodge DM, Ruiz GM, Maclsaac HJ (2007) Predicting invasion risk using measures of introduction effort and environmental niche models. *Ecological Applications*, **17**, 663-674
- Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773-785
- Hidalgo O, Garcia-Jacas N, Garnatje T, Susanna A (2006) Phylogeny of Rhaponticum (Asteraceae, Cardueae–Centaureinae) and related genera inferred from nuclear and chloroplast DNA sequence data: taxonomic and biogeographic implications. *Annals of Botany*, **97**, 705-714
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965-1978
- Hutchinson GE (1957) Concluding remarks. *Cold Spring Harbor Symposium Quantitative Biology*, **22**, 415– 427.
- Ibáñez I, Silander Jr JA, Wilson AM, LaFleur N, Tanaka N, Tsuyama I (2009) Multivariate forecasts of potential distributions of invasive plant species. *Ecological Applications*, **19**, 359-375
- Jiménez-Valverde A, Lobo JM (2007) Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecologica*, **31**, 361-369
- Jiménez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, **14**, 885-890
- Jiménez-Valverde A, Peterson A, Soberón J, Overton J, Aragón P, Lobo J (2011) Use of niche models in invasive species risk assessments. *Biological Invasions*, **13**, 2785-2797
- Koloren, O., S. Uygur, O. Bozdogan, F.N. Uygur & U. Schaffner 2008. Population density and reproductive output of *Acroptilon repens* L., in Turkey. *Pakistan Journal of Botany*, **40** , 2259-2263.

- Lavergne S, Muenke NJ, Molofsky J (2010) Genome size reduction can trigger rapid phenotypic evolution in invasive plants. *Annals of Botany*, **105**,109-116
- Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385-393
- López Alvarado J, CrespoVillalba MB, García-Jacas N, et al. (2011) First record of the alien pest *Rhaponticum repens* (Compositae) in the Iberian Peninsula. *Collectanea Botanica*, **30**, 59-62
- López-Darias M, Lobo JM (2008) Factors affecting invasive species abundance: the Barbary ground squirrel on Fuerteventura Island, Spain. *Zoological Studies*, **47**, 268-281
- Maddox, Donald M.; Mayfield, Aubrey; Poritz, Noah H. 1985. Distribution of yellow starthistle (*Centaurea solstitialis*) and Russian knapweed (*Centaurea repens*). *Weed Science*, **33**, 315-327.
- da Mata RA, Tidon R, Côrtes LG, De Marco P, Diniz-Filho JAF (2010) Invasive and flexible: niche shift in the drosophilid *Zaprionus indianus* (Insecta, Diptera). *Biological Invasions*, **12**, 1231-1241
- Medley KA (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
- Millenium Ecosystem Assessment (2005) Millennium ecosystem assessment. Ecosystems and Human Well-being: Synthesis Washington DC: Island Press
- Mukherjee A, Williams D, Wheeler G, Cuda J, Pal S, Overholt W (2011) Brazilian peppertree (*Schinus molle*) in Florida and South America: evidence of a possible niche shift driven by hybridization. *Biological Invasions*. In press.
- Nenzén H, Araújo M (2011) Choice of threshold alters projections of species range shifts under climate change. *Ecological Modeling*, **222**, 3346-3354
- Olson DM, Dinerstein E, Wikramanayake ED, et al. (2001) Terrestrial ecoregions of the world: a new map of life on earth. *Bioscience*, **51**, 933-938.
- Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology*, **78**, 419-433

- Peterson AT (2005) Predicting potential geographic distributions of invading species. *Current Science*, **89**, 9
- Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102-117
- 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 SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231-259
- Phillips BL, Chipperfield JD, Kearney MR (2008) The toad ahead: challenges of modelling the range and spread of an invasive species. *Wildlife Research*, **35**, 222-234.
- 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.
- Richardson DM, Thuiller W (2007) Home away from home—objective mapping of high-risk source areas for plant introductions. *Diversity and Distributions*, **13**, 299-312
- Roura-Pascual N, Suarez AV, Gómez C, et al. (2004) Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **271**, 2527-2535
- Roura-Pascual N, Bas JM, Thuiller W, Hui C, Krug RM, Brotons L (2009) From introduction to equilibrium: reconstructing the invasive pathways of the Argentine ant in a Mediterranean region. *Global Change Biology*, **15**, 2101-2115
- Santos M, Freitas R, Crespí AL, Hughes SJ, Cabral JA (2011) Predicting trends of invasive plants richness using local socio-economic data: An application in North Portugal. *Environmental Research*, **111**, 960-966
- Schierenbeck KA, Ellstrand NC (2009) Hybridization and the evolution of invasiveness in plants and other organisms. *Biological Invasions*, **11**, 1093-1105
- Schoener TW (1968) The *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology*, **49**, 704-726

- Stevens, K.L., 1986. Allelopathic polyacetylenes from *Centaurea repens* (Russian knapweed). *Journal of Chemical Ecology*, **12**, 1205-1211
- Soberón J (2007) Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters*, **10**, 1115-1123
- de Souza Muñoz ME, De Giovanni R, de Siqueira MF, et al. (2011) openModeller: a generic approach to species' potential distribution modelling. *Geoinformatica*, **15**, 111-135
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285-1293
- Urban MC, Phillips BL, Skelly DK, Shine R (2007) The cane toad's (*Chaunus* [Bufo] *marinus*) increasing ability to invade Australia is revealed by a dynamically updated range model. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 1413-1419
- Václavík T, Meentemeyer RK (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modeling*, **220**, 3248-3258
- Van der Vaart AW (2000) Asymptotic statistics. Cambridge Univ Press
- Warren DL, Glor RE, Turelli M (2008) Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution*, **62**, 2868-2883
- Warren DL, Seifert SN (2011) Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological Applications*, **21**, 335-342
- Watson, A. K. 1980. The biology of Canadian weeds. 43. *Acroptilon* (*Centaurea*) *repens* (L.) DC. *Canadian Journal of Plant Science*, **60**, 993-1004,
- Wiens JJ, Ackerly DD, Allen AP, et al. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310-1324
- Wisz MS, Hijmans R, Li J, Peterson AT, Graham CH, Guisan A (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763-773

7. ON CONSIDERING THE TEMPORAL DIMENSION

STUDY: ‘Species-specific exposure to climate change in time and space: from climate velocity to bioclimatic-velocity’

7.1 Introduction

Assessing vulnerability of terrestrial ecosystems to climate change over the next 50-100 years is a highly uncertain and complex task. This complexity is illustrated through divergent outcomes projected by a variety of methodological approaches. One of the most widespread techniques used to inform conservation and 21st century climate adaptation is species distribution modeling (SDM) (e.g. Thomas et al., 2004; Thuiller, 2004; Araújo et al., 2011).

SDM correlates species presence or abundance to climate and other environmental variables, typically using statistical learning methods, so the bioclimatic profile of the species is quantified (Franklin, 2010). Subsequently, this climatic profile is applied to mapped climate projections, to evaluate which areas will be more or less suitable relative to present conditions. It has been recognized that this approach is limited to the assessment of exposure to climate change (extent of climate change likely to be experienced by a species; after Dawson et al. (2011)). Species traits such as ecophysiological or life history sensitivity (Woodward, 1992; Keenan et al., 2011) and adaptive capacity (Davis & Shaw, 2001; Sork et al., 2010) may well constrain and even contradict results from only climate change exposure analysis. Nevertheless, the benefit of assessing exposure to climate change is that it may be estimated without very detailed information on the target species and is useful for bounding the range of actions decision makers may need to consider (see Figure 3 in Dawson et al., 2011).

Recent research has focused on developing methods to account for how quickly climatic changes are occurring across different areas (Loarie et al., 2009; Ackerly et al., 2010; Burrows et al., 2011). For instance, Loarie et al. (2009) derived climate velocity ($\text{km}\cdot\text{yr}^{-1}$) as the ratio of absolute projected local temperature change per year over the spatial gradient in projected temperature. They used the measure to examine patterns of exposure and conservation risk for the world's major biomes. Ackerly et al. (2010) mapped and analyzed local climate change velocity in California to help identify the magnitude and pattern of biodiversity risk. Moreover, the velocity of past climatic changes have resulted in clear effects on species extinction and evolution (Nogués-Bravo et al., 2010; Sandel et al., 2011), therefore it is important to detect high velocities under the rapid ongoing climatic warming, as species' capacities for adaptation and migration may be challenged (Davis & Shaw, 2001).

Presumably, climate velocity is proportional to the rate at which the biota of an area must migrate locally in order to encounter or track shifts in its suitable habitat according to its correlation with recent historical climate. However, as Ackerly et al. (2010) point out, ecologists anticipate species to manifest distinct, individualistic responses to climate change. Research is needed to better understand how species climate change exposures could vary as a function of the local climate change velocities.

We investigated the velocity of species exposure to predicted climate changes using a consensus measurement of species habitat suitability derived from general climatic variables. We studied time-exposure dynamics of endemic oak and pine trees in a Mediterranean climate region (California Floristic Province), which has been determined to be one of the most sensitive biomes to climate change globally (Sala & Chapin, 2000; Underwood et al., 2009). Oaks and pines were selected because they are diverse and widespread ecologically (Pavlik et al., 1991; Richardson, 2000). Specifically, we analyzed spatio-temporal dynamics in species potential suitable area and current distribution of stands, based on species distribution models for recent historical climate (1971-2000) and projected climates for mid-century (2041-2070) and end of century (2071-2100).

7.2 Material and methods

Species distribution models (SDMs) were developed for eight ecologically widespread species of oak and pine trees endemic to the California Floristic Province: *Pinus balfourniana*, *P. coulteri*, *P. muricata*, *P. sabiniana*, *Quercus douglasii*, *Q. engelmannii*, *Q. lobata*, and *Q. wislizenii*. Species presence data from recent vegetation survey plots were used to model species occurrence probabilities as a function of mapped bioclimatic variables.

Species occurrence data were extracted from 42 existing vegetation inventories (compiled by Hannah et al., 2008). Only records from presence/absence vegetation surveys were used. The following current climatic variables were obtained from a statistical downscaling (Flint & Flint, 2012) of the PRISM database (PRISM Climate Group, Oregon State University, available at: <http://prism.oregonstate.edu>) from 800 m to 270 m spatial resolution: total annual precipitation, mean annual temperature, precipitation of the driest month, maximum temperature of the warmest month and minimum temperature of the coldest month. Although different species may be limited more or less by different subsets of these variables, we included all variables in the models to ensure maximum comparability among species.

Using averaged climate data for the period 1971-2000, we estimated eight different SDMs for each species within the BIOMOD platform (Thuiller et al., 2009 and see Appendix 4.1 for model descriptions) in order to obtain a robust measure of species climatic suitability (Araújo & New, 2007). Model calibration was undertaken with 70% of the presence-absence observations and the remaining 30% were used for validation. We applied a random selection of absences equal to two times the number of presences for each species, allowing for a large number of absences covering the environmental space (Barbet-Massin et al., 2012). In addition, each model was run twice using different random samples of absence to address sample bias in absences (Elith et al., 2010). Models were able to reproduce current distributions with good accuracy (Appendix 4.2).

In order to discriminate suitable from unsuitable areas, a threshold in the continuous climatic suitability measure predicted by the SDM was applied to each model. We set the threshold by maximizing the True Skill Statistic (TSS) metric (Allouche et al., 2006). Estimates of current and

future suitable area were then identified based on the agreement of at least 5 models in considering an area suitable after threshold binarization. A continuous consensus climatic suitability was then obtained for each grid cell by averaging the probabilities from those models agreeing with the consensus of suitable area.

Habitat suitability dynamics were estimated based on differences between models' consensus projection for current climate and for future projected climate under the A2 emissions scenario using the GFDL (Geophysical Fluid Dynamics Laboratory) global circulation model (GCM). This GCM is one of two models that were selected for climate change assessment in California (the other being the NCAR Parallel Climate Model (PCM)) because of its ability to reproduce historic climate patterns accurately (Cayan et al., 2008). This combination of GCM and scenario represents a "strong change" scenario of a much warmer and drier California used by the California Climate Change Center (Cayan, 2009) for impact analysis and to derive informed conservation policies. Two climate change periods were projected: mid-21st century (averaged climate 2041-2070; period 1) and the late-21st century (averaged climate 2071-2100; period 2).

We computed the bio-velocity of climate change using the same procedure as in Loarie et al. (2009) but applied it to each species' suitability map rather than to a single climate variable. We divided the temporal gradient (e.g. magnitude of change over time) of climate suitability by the spatial gradient (e.g. magnitude of change over space) in suitability for the period under analysis. Temporal gradient is computed as the difference in consensus probabilities between present and future projection per unit of time (years): 70 years for period 1 and 100 years for period 2. Spatial gradients are computed as the slope of probabilities using the maximum average technique (Burrough et al., 1998) in a 9-pixel kernel. To avoid infinite velocities, we excluded flat spatial gradients (< 0.001). The result is a velocity measure of the changes in climatic suitability of each species.

For each of the two time periods, we calculated several metrics to evaluate species' exposure to predicted climate change at two different organizational scales: range level (climatically suitable areas) and plot level (current forests). Range-level metrics describing species exposure provide information most relevant to developing conservation strategies that address broad patterns of

change in climate suitability, including potential new areas for colonization (whether assisted or not), whereas the plot-level metrics inform more local management strategies focused on species' adaptation and in situ conservation (see Appendix 4.3 for full description of the metrics).

At the range level, five metrics were calculated:

(1) species range change (SRC), which measures differences of potential suitable area (Thuiller et al., 2005) per year and is related to exposure to extinction;

(2) range exposure to migration (REM), calculated as the difference in suitable habitat area between full and null dispersal assumptions (Svenning & Skov, 2004; Araújo & New, 2007) divided by the time lapse between the current and targeted period, which emphasizes the relevance of migration processes in lowering exposure;

(3) range change velocity (RCV), calculated as the net balance between trailing edge and leading edge velocity based on the average of climatic suitability by grid cell of models coincident with consensus range dynamics, which identifies potential disruptions in the edges of ranges to tracking climate change;

(4) range spatial fragmentation (RSF), calculated as the number of discrete habitat patches (McGarigal 2006);

(5) range spatial aggregation (RSA), calculated as the percent of total suitable habitat occupied by the largest patch (McGarigal, 2006). We assessed the temporal rates of change in RSF and RSA by dividing current and future projections of the metric by the time elapsed between projections. These landscape metrics assess the spatial configuration of potential suitable habitat, which is related to population persistence (Opdam & Wascher, 2004).

To assess plot-level exposure of current forests, we defined three metrics:

(1) forest migration effort (FME), which measures the mean distance of forest plots to projected future climatically suitable area using a least cost-distance route based on suitability measures. Skov & Svenning (2004) used a similar approach based on tree cover to assess potential migration routes for European herbs, and Wang et al. (2008) found a significant relationship between gene flow and a suitability resistance measure;

(2) forest climate-site exposure (FCE), calculated as the percentage of forest plot locations switching from suitable to unsuitable conditions based on the set threshold of habitat suitability.

Although this measure has been used as a surrogate for extinction risk (Thomas et al., 2004), we have adopted it as a forest exposure risk to new climatic conditions, and

(3) forest climatic velocity (FCV), the mean bioclimatic velocity of forests plots decreasing in suitability, calculated by overlaying forest plots with the bioclimatic velocity grid computed using the methods described above.

7.3 Results

Bioclimatic velocities differ greatly among species and climate change projections in the periods analyzed (Figure 7.1). In some cases, current distributional area is located at the transition between slowly increasing and decreasing predicted bioclimatic velocities (e.g., Fig 7.1a, *P. sabiniana*), whereas in other cases, current climatic suitable area is located where bioclimatic suitability is predicted to rapidly decrease (e.g., *P. balfourniana*). However, for these California mountain pines and oaks, high velocities leading to climatic unsuitability are located in and around the Central Valley, whereas high velocities increasing climatic suitability tend to concentrate in Northern mountain ranges, although this pattern is quite species-idiosyncratic. Predicted velocities are also markedly different for each period of analysis, both in extent and spatial distribution (Figure 7.1b). For instance, for *Q. douglasii* and *P. coulteri* larger areas are predicted to experience higher velocities in period 1, whereas higher velocities are predicted to occur in period 2 for *Q. lobata* and *P. muricata*.

Species ranges (the extent of climatically suitable habitats) are predicted to shrink at greater rates in period 1 (species range change, SRC), for all species except *P. muricata*, *Q. lobata* and *Q. wislizenii* (Figure 7.2a). Climatically suitable habitat for *Q. engelmannii* is predicted to shrink at a rate of over 1.2 % per year. On the other hand, *P. sabiniana* is predicted to expand its climatically suitable range and the projected rate tends to increase with time. Because SRC does not account for dispersal, these rates of range dynamics should be interpreted together with the information provided by range exposure to migration (REM, Figure 7.2b). REM represents the extent to which the role of dispersal potentially becomes relevant by examining the ratio between full dispersal and null dispersal assumptions in species ranges (Guisan & Thuiller 2005). Our predictions suggest that

REM will generally be higher in period 2, except for *P. balfourniana* and *Q. douglasii*, where high rates of projected habitat loss tend to diminish differences between full dispersal and null dispersal assumptions.

Differences in predicted velocities between trailing edge and leading edge (range change velocities, RCV) highlight the dynamism and heterogeneity of climatically suitable habitat for each species and period (Figure 7.2c and 7.2d). In general, velocity in ranges tends to be higher in period 2 except for *Q. douglasii* and *P. balfourniana* leading edge and *P. coulterii* and *Q. engelmannii* trailing edge (Figure 7.2c). Leading edge velocities are higher than trailing edge velocities for the majority of species and periods (Figure 7.2d). Interestingly, two species present a reversal in their response across periods: *P. balfourniana* presents higher leading edge velocities only in period 1 whereas *Q. douglasii* in period 2. Net differences between trailing and leading edge follow a general pattern of increase in velocity in period 2, according to the results in each edge. Nevertheless, differences in the magnitude are noteworthy and range from very large increase (e.g. *P. sabiniana*) to low increase (e.g. *P. muricata*).

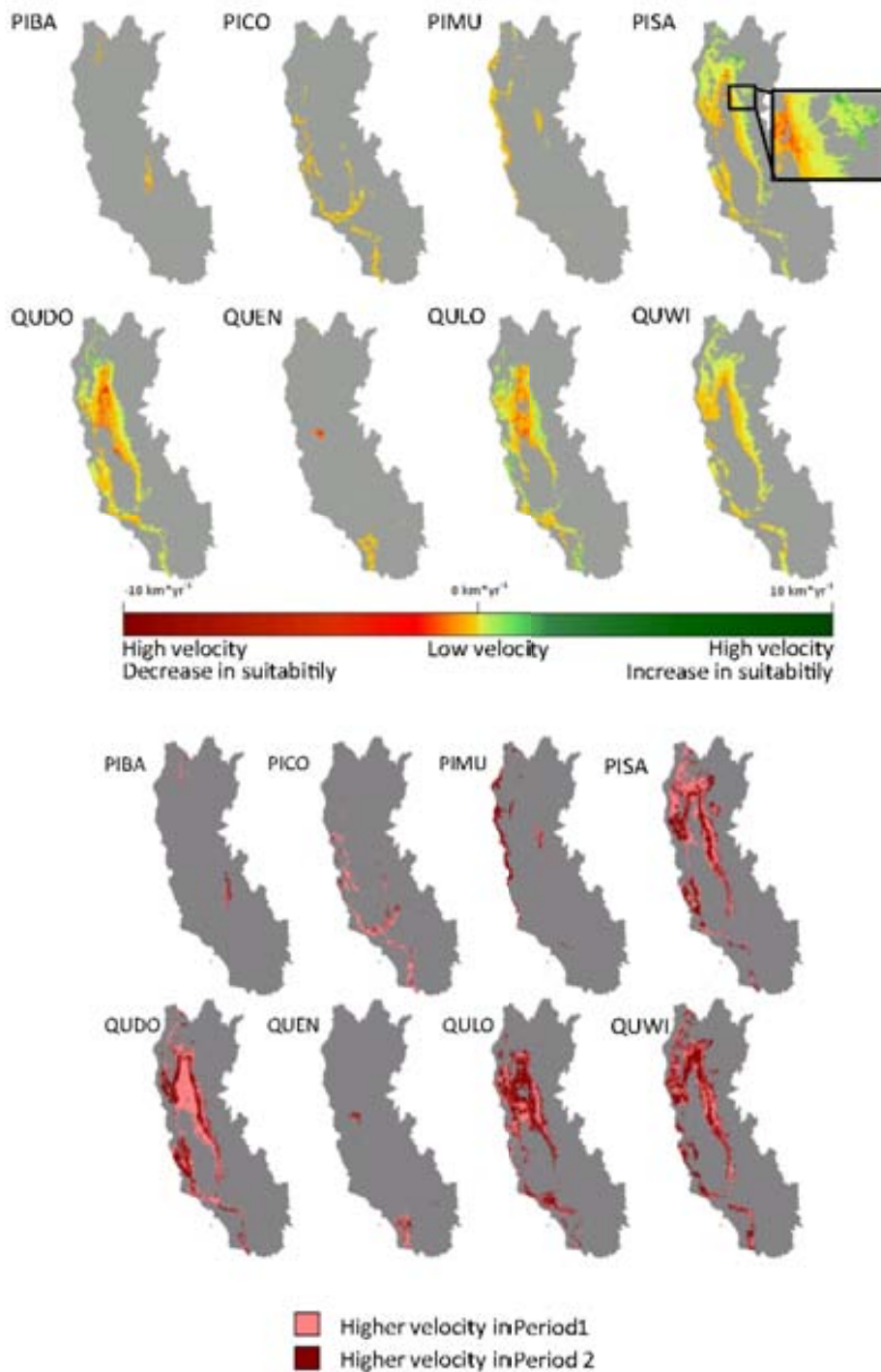


Figure 7.1 The bioclimatic velocity of climate change for different California endemic tree species within their range. (a) Bioclimatic velocity for the period present (1971-2000) to mid-century (2041-2070). (b) Temporal categorization bioclimatic velocity, whether higher in mid-century (2041-2070) or end of century (2071-2100). Gray color indicates lack of spatial gradient or not suitable areas in any time frame. PIBA = *P. balfourniana*, PICO = *P. coulteri*, PIMU = *P. muricata*, PISA = *P. sabiniana*, QUEN = *Q. engelmannii*, QUDO = *Q. douglasii*, QULO = *Q. lobata*, QUWI = *Q. wislizenii*.

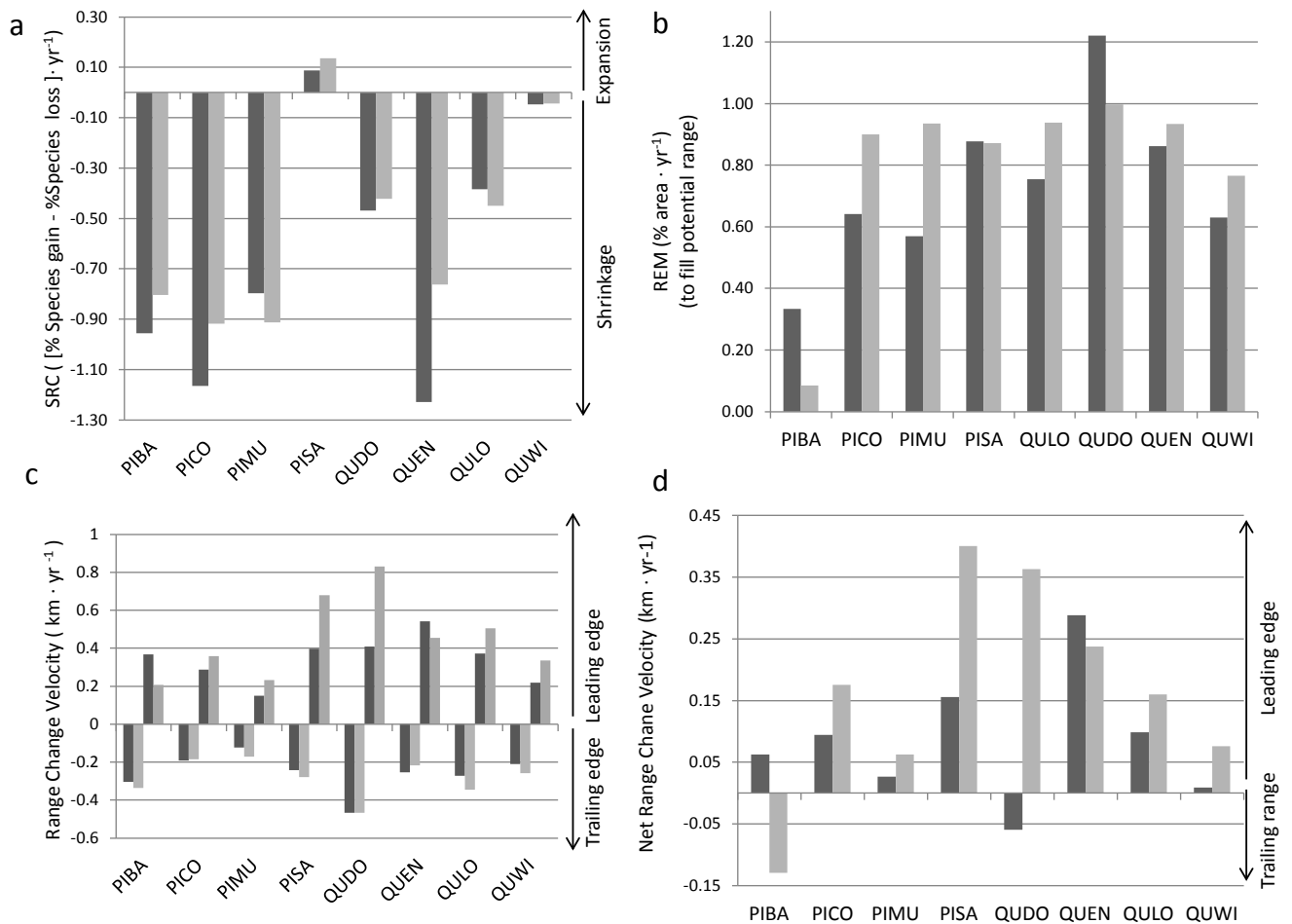


Figure 7.2 Climate change exposure in species ranges in to period 1 (present (1971-2000) to mid-century (2041-2070), dark grey) and period 2 (present (1971-2000) to end of century (2041-2070), light grey). (a) Species range change (SRC). (b) Range exposure to movement (REM): time rates between full versus null dispersal in their ranges. (c) Species range velocity in trailing versus leading ranges: “Negative” velocities indicate loss of climatic suitability (trailing edge) whereas “positive” velocities indicate an increase of climatic suitability (leading edge). (d) Range change velocity (RCV): differences between velocities of leading edge and trailing edge. Species abbreviations defined in Figure 1 caption

Spatial configuration of predicted habitat highlights different velocities in aggregation (range spatial aggregation, or RSA) and fragmentation processes (range spatial fragmentation, or RSF) (Figure 7.3) of climatically suitable habitat. Together with shrinkage dynamics, yearly proportional declines in the number of suitable patches per year (a measure of RSF) and the percentage of the largest patch area (describing RSA) are projected for several species: *P. balfourniana*, *P. coulteri* and *P. muricata*, although the rates of such changes vary between periods (Figure 7.3a and 7.3b).

It is worth noting some additional projected patterns: *Q. engelmannii* shows a predicted pattern of fragmentation and decline, with an increasing rate of the number of patches between period 1 to period 2, and decline in the area of the largest patch. A similar pattern is predicted for *Q. douglasii*, although period 1 shows higher rates of fragmentation than period 2. Another notable response occurs in *Q. wislizenii*, with increasing fragmentation predicted in the first period and aggregation (negative rates in the number of patches but increase in the largest patch area) in the second. Overall, some predicted patterns of fragmentation and aggregation of suitable habitat may be identified, but our results demonstrate that the velocity in spatial configuration varies in a heterogeneous manner and is species-specific.

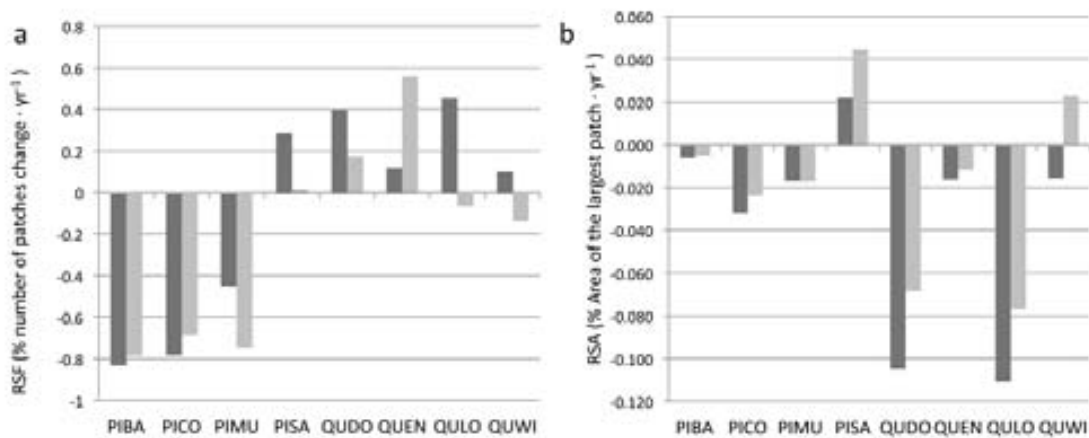


Figure 7.3 Spatial configuration dynamics in species ranges in period 1 (present (1971-2000) to mid-century (2041-2070), dark grey) and period 2 (present (1971-2000) to end of century (2041-2070), light grey). (a) Range spatial fragmentation (RSF). Percentage of patch abundance increase/decrease per unit of time. (b) Range spatial aggregation (RSA). Percentage of area change of the species' largest suitable habitat patch per unit of time. Species abbreviations defined in Figure 1 caption.

Projected exposure of current forest distribution suggests different dynamics than for projected future ranges. Current forest locations occupied by each species present different degrees of velocity change between periods ranging from subtle increases in period 2 (*P. balfourniana*, *P. sabiniana* and *Q. douglasii*) to much larger velocities in period 1 (*P. coulterii* and *Q. lobata*) (Figure 7.4a). These bioclimatic velocities in plots (FCV) should be reflected in the percentage of plots becoming unsuitable per unit of time (forest change in suitability, or FCS) (Figure 7.4b), but this is not always the case. For instance, for *P. balfourniana* FCV is higher in period 2 whereas FCS is higher

in period 1. Such situations imply a nuanced difference between the two metrics: while FCV measures the velocity of the gradient, FCS measures the rate after which a certain threshold is achieved and forests become exposed. We acknowledge the uncertainty deriving from the method used to identify this threshold and its consequences in climate change projections (Nenzén & Araújo 2011). Nevertheless, distinguishing the two measures is useful because each has different implications relative to conservation strategies

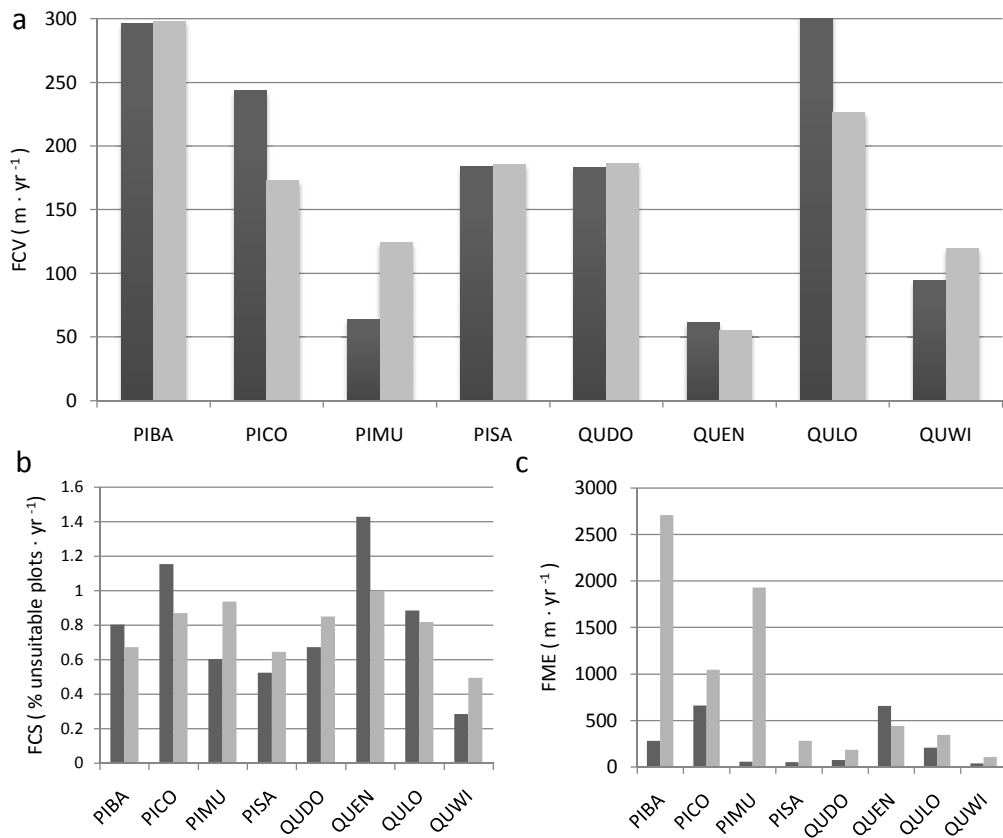


Figure 7.4 Velocity of climate change exposure on current distribution in period 1 (present (1971-2000) to mid-century (2041-2070), dark grey) and period 2 (present (1971-2000) to end of century (2041-2070), light grey). (a) Forest velocity change (FVC): average velocity of bioclimatic exposure in species plots. (b) Forest change suitability (FCS): percentage of plots becoming unsuitable per unit of time (c) Forest migration exposure (FME): average of cost-distance to the nearest suitable patch. Species' abbreviations defined in Figure 1 caption.

The projected rate of climatic isolation of current forests, or forest migration exposure (FME, Figure 7.4c), as measured by the average least cost-distance to the nearest suitable patch, leads to a large increase in isolation during period 2, especially for *P. balfourniana*, *P. coulteri* and *P. muricata*. The exception of *Q. engelmannii*. which is caused by the predicted total loss of suitable habitat in the second period. Clearly, *P. balfourniana*, *P. coulteri* and *P. muricata* migration (whether assisted or not) will need further consideration since these species are projected to be exposed to an accelerating source of risk through the following decades, requiring predicted migration rates of up to 2712, 1930 and 1047 m·yr⁻¹ respectively to reach suitable habitat. The case of *P.balfourniana* is especially indicative of two disjunct populations, one of them becoming unsuitable in the second period, which substantially increases the FME indicator.

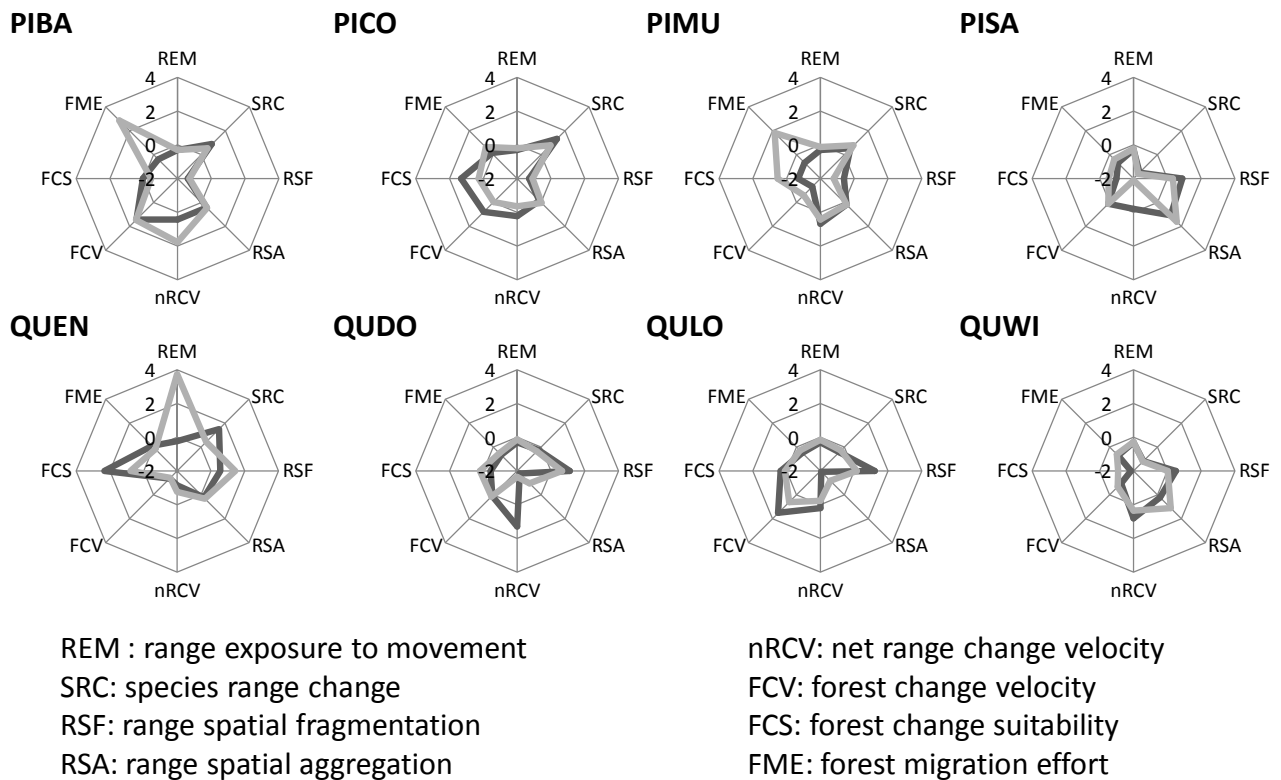


Figure 7.5 Species' exposure to climate change for different dimensions and periods (period 1 -present (1971-2000) to mid-century (2041-2070)-, dark grey; period 2 -present (1971-2000) to end of century (2041-2070)-, light grey). Data have been z-transformed for each variable and period in order to compare species and dimensions. Species abbreviations defined in Figure 1 caption.

Comparing different species and exposure dimensions reveals the complexity and dynamism of exposure to climate change, which is contingent on the species and varies in time (Figure 7.5). While for some species there is a similarity between exposure dimensions for period 1 and 2 (*Q. lobata*, *P. coulterii*), for others (*P. balfourniana*, *P. muricata*, *Q. engelmannii*) exposure is dramatically increased or decreased in one or more dimensions, pointing to several processes (e.g. dispersal, migration, etc.) that may become key sources of exposure in each time frame analyzed. For example, for *P. balfourniana* and *P. muricata* current forests may be challenged by the requirements for dispersal, whereas *Q. douglasii* will have to cope with rapidly varying climatic conditions in period 1. Additionally, this graphical summary emphasizes differences between exposures at the range level versus exposures at the current distribution level.

7.4 Discussion

We have shown that the projected bioclimatic velocities of species' exposure to climate change vary widely depending on the species under analysis. Our results illustrate that bio-velocity and temporal measurements applied to spatial analysis of species' distributions can reveal rather complex interacting dynamics in the form of differences between leading and trailing edge velocities, varied habitat fragmentation and aggregation patterns, migration challenges and differences in in-situ forest exposure velocity. Our results suggest that targeted conservation responses will be required sooner rather than later for some endemic oaks and pines of the California Floristic Province as some key processes are predicted to be challenged more rapidly from now to mid-century than from mid-century to the end of the century. However this will be highly depending on the exposure dimension and species under consideration.

In general, most species analyzed here tend to show higher velocities in their range dynamics by mid- century, but current forests exposure and net differences between velocities in trailing and leading edges draw a complex picture of the pace of climate suitability increasing and decreasing in some areas. Accordingly, several modeling and observational studies have also identified diverse patterns of range dynamics. Some studies have shown that the leading edge of the range may become occupied at higher rates than trailing edge (Chen et al. 2011 and references therein) and

accumulation of extinction debt may occur at trailing edges especially for long-lived organism, such as trees (Kuussaari et al., 2009). These effects could eventually result in range expansions. On the other hand, empirical observations from eastern US forests show that the common current signal is range erosion (Murphy et al., 2010). In line with this, a recent analysis of coupled species distribution and population models found high rates of range reduction (44% approximately) for alpine plants (Dullinger et al., 2012), similar to projections of niche based models under an unlimited dispersal scenario. In any case, we acknowledge that bioclimatic velocity may over- or under-emphasize range erosion unless it is not balanced with explicit spatial and temporal population dynamics.

It is likely that temporal resolution (30 year averaged climate in this analysis) may affect the results found here. Indeed extreme climatic events, or climate change at a finer temporal resolution, may accelerate species' responses and produce non-linearities not reported in our study (Easterling et al., 2000; Coumou & Rahmstorf, 2012). For instance, extreme climatic events may be important for migration routes for some taxa (Early & Sax, 2011), large-scale vegetation responses and food webs (Carnicer et al., 2011). However, there is also potential for vegetation resilience to extreme climatic events making true distributional shifts in forests the exception rather than the rule (Lloret et al., 2012). Considering the wide range of species-specific outcomes that may occur within short time frames, we argue that species exposure predicted at a 30-year temporal resolution provides a useful climate risk context in which to embed predictions of extreme climatic events. Future research could, for instance, detect non-stationary dynamics and tipping-points in species exposure to new conditions (e.g. climate change) using yearly sliding-window analysis through the entire time sequence analyzed (2041-2100).

Species' exposure to environmental changes is multidimensional and also time dependent. We anticipate that the different paces of changing conditions will yield potentially different synergies between exposure dimensions. For example, a species may be characterized by high velocity leading to exposure of current forests and increased migration effort in period 1, whereas period 2 presents a decrease in habitat area and increased fragmentation that may constitute a significant threat. Altogether, such varying exposures in time should be taken into account in conservation planning to define actions and priorities in an appropriate time frame (Hannah et al., 2002;

Mawdsley et al., 2009). Furthermore, the pace of changing climatic exposure suggests the inclusion of more dynamic parameters of habitat suitability in new hybrid modeling approaches dealing with the challenge to predict species distributions (Dormann et al., 2012).

The distinction between two complementary levels (plot and range level) is important in exposure risk assessment. Range level, measuring areas of potential suitable future conditions, may over- or under- represent risk because SDM have errors of commission and omission not detected in performance metrics (Lobo et al., 2008). In contrast, plot-level depicts exposure at current locations where the species is actually observed, therefore it provides a forecast of the current forests under risk. It is noteworthy that the difference between the two measures constitutes an interesting approach to global change projections: higher differences between range and plot levels could indicate that vulnerability may be either be more dependent on in-situ adaptive capacity (higher plot level exposure) or more dependent on migration and colonization capacity (higher range-level exposure). Such differences are also likely to be time dependent, therefore calling for a dynamic conservation strategy (Hannah & Hansen, 2005).

The results presented in this study show that in principle, climate velocity maps should be species-specific. In practice, we acknowledge that challenges presented by the large number of species and the limited information on climate associations for many species makes it difficult to analyze and synthesize climate velocity species-by-species. Furthermore, it is also challenging to develop unified conservation and forest management strategies for species responding quite differently to changing conditions in space and time. We argue that, when possible and for endangered species or species that provide ecosystem services, species-specific velocity maps provide new valuable information on the dynamics of species exposures to climate change in the coming century, helping to identify conservation action priorities taking into account, also, the temporal scale.

7.4 References

- Ackerly D, Loarie S, Cornwell W et al. (2010) The geography of climate change: Implications for conservation biogeography. *Diversity and Distributions*, **16**, 476-487.
- 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 MB, Alagador D, Cabeza M, Nogués-Bravo D, Thuiller W (2011) Climate change threatens European conservation areas. *Ecology Letters*, **14**, 484-492.
- Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends in Ecology and Evolution*, **22**, 42-47.
- Barbet-Massin M, Jiguet F, Albert CH, Thuiller W (2012) Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, **3**, 327-338.
- Burrough PA, McDonnell RA (1998) Principles of GIS. Oxford Univ.Press.
- Burrows MT, Schoeman D S, Buckley LB et al. (2011) The pace of shifting climate in marine and terrestrial ecosystems. *Science*, **334**, 652-655.
- Carnicer J, Coll M, Ninyerola M et al. (2011) Widespread crown condition decline, food web disruption, and amplified tree mortality with increased climate change-type drought. *Proceedings of the National Academy of Science of the U S A*, **108**, 1474.
- Cayan DR, Maurer EP, Dettinger MD, Tyree M, Hayhoe K (2008) Climate change scenarios for the California region. *Climatic Change*, **87**, 21-42.
- Cayan DR, Tyree M, Dettinger MD et al. (2009) Climate change scenarios and sea level rise estimates for the California 2009 Climate Change scenario Assessment. California Energy Commission-California Ocean Protection Council-California Environmental Protection Agency.
- Chen I, Hill JK, Shiu H et al. (2011) Asymmetric boundary shifts of tropical montane lepidoptera over four decades of climate warming. *Global Ecology and Biogeography*, **20**, 34-45.

- Coumou D, Rahmstorf S (2012) A decade of weather extremes. *Nature Climate Change*, In press.
- Davis M, Shaw R(2001) Range shifts and adaptive responses to Quaternary climate change. *Science*, **292**, 673-679.
- Dawson TP, Jackson ST, House JI, Prentice IC, Mace GM (2011) Beyond predictions: Biodiversity conservation in a changing climate. *Science*, **332**, 53-58.
- Dormann CF, Schymanski SJ, Cabral J et al. (2012) Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*. In press.
- Dullinger S, Gatttringer A, Thuiller W et al. (2012). Extinction debt of high-mountain plants under twenty-first-century climate change. *Nature Climate Change*. In press.
- Early R, Sax DF (2011) Analysis of climate paths reveals potential limitations on species range shifts. *Ecology Letters*, **14**, 1125-1133.
- Easterling DR, Meehl GA, Parmesan C et al. (2000) Climate extremes: Observations, modeling, and impacts. *Science*, **289**, 2068-2074.
- Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. *Methods in Ecology and Evolution*, **1**, 330-342.
- Flint L, Flint A (2012) Downscaling future climatic scenarios to fine scales for hydrologic and ecological modeling and analysis. *Ecological Processes*, **1**, 2.
- Franklin J (2010) Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press, Cambridge, UK
- Guisan A, Thuiller W (2005) Predicting species distribution: Offering more than simple habitat models. *Ecology Letters*, **8**, 993-1009.
- Hannah L, Midgley G, Davies I et al. (2008) BioMove - Improvement and Parameterization of a Hybrid Model for the Assessment of Climate Change Impacts on the Vegetation of California. CEC-500-02-004, California Energy Commission, Public Interest Energy Research Program.
- Hannah L, Midgley GF, Lovejoy T et al. (2002) Conservation of biodiversity in a changing climate. *Conservation Biology*, **16**, 264-268.

- Hannah L, Hansen L (2005) Designing landscapes and seascapes for change. Pages 329–341 in T. E. Lovejoy and L. Hannah, editors. *Climate Change and Biodiversity*. Yale University Press, New Haven, Connecticut.
- Keenan T, Serra JM, Lloret F, Ninyerola M, Sabaté S (2011) Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO₂ matters!. *Global Change Biology*, **17**, 565-579.
- Kuussaari M, Bommarco R, Heikkinen RK et al. (2009). Extinction debt: A challenge for biodiversity conservation. *Trends in Ecology and Evolution*, **24**, 564-571
- Lloret F, Escudero A, Iriondo JM, Martínez-Vilalta J, Valladares F (2012) Extreme climatic events and vegetation: The role of stabilizing processes. *Global Change Biology*, **18**, 797-805.
- Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, Ackerly DD (2009) The velocity of climate change. *Nature*, **462**, 1052-1055.
- Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.
- Mawdsley JR, O'Malley R, Ojima DS (2009) A review of climate-change adaptation strategies for wildlife management and biodiversity conservation. *Conservation Biology*, **23**, 1080-1089.
- McGarigal K (2006) Landscape Pattern Metrics. *Encyclopedia of Environmetrics*. John Wiley & Sons, Ltd.
- Murphy HT, VanDerWal J, Lovett-Doust J (2010) Signatures of range expansion and erosion in eastern North American trees. *Ecology Letters*, **13**, 1233-1244.
- Nenzén H, Araújo M (2011) Choice of threshold alters projections of species range shifts under climate change. *Ecological Modelling*, **222**, 3346-3354.
- Nogués-Bravo D, Ohlemüller R, Batra P, Araújo MB (2010). Climate predictors of late quaternary extinctions. *Evolution*, **64**, 2442-2449.
- Opdam P, Wascher D (2004) Climate change meets habitat fragmentation: Linking landscape and biogeographical scale levels in research and conservation. *Biological Conservation*, **117**, 285-297

- Pavlik BM, Muick PC, Johnson S, Popper M (1991) Oaks of California. Cachuma Press, Los Olivos, CA. ISBN: 0-9628505-1-9
- Richardson DM (2000) Ecology and Biogeography of Pinus. Cambridge Univ Pr, Cambridge, UK
- Sala OE, Chapin F (2000) Global biodiversity scenarios for the year 2100. *Science*, **287**, 1770.
- Sandel B, Arge L, Dalsgaard B et al. (2011) The influence of late Quaternary climate-change velocity on species endemism. *Science*, **334**, 660-664.
- Skov F, Svenning J (2004) Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography*, **27**, 366-380.
- Sork VL, Davis FW, Westfall R et al. (2010) Gene movement and genetic association with regional climate gradients in California Valley Oak (*Quercus lobata* née) in the face of climate change. *Molecular Ecology*, **19**, 3806-3823.
- Svenning J, Skov F (2004) Limited filling of the potential range in European tree species. *Ecology Letters*, **7**, 565-573.
- Thomas CD, Cameron A, Green RE et al. (2004) Extinction risk from climate change. *Nature*, **427**, 145-148.
- Thuiller W (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020-2027.
- Thuiller W, Lavorel S, Araújo MB, Sykes MT, Prentice IC (2005) Climate change threats to plant diversity in Europe. *Proceeding National Academy Science U S A*, **102**, 8245-8250.
- Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD : A platform for ensemble forecasting of species distributions. *Ecography*, **32**, 369-373.
- Underwood EC, Viers JH, Klausmeyer KR, Cox RL, Shaw MR (2009) Threats and biodiversity in the mediterranean biome. *Diversity and Distributions*, **15**, 188-197.
- Wang YH, Yang KC, Bridgman CL, Lin LK (2008) Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecology*, **23**, 989-1000.

8. FINAL REMARKS AND LESSONS LEARNED

8.1 Reflecting on the use of correlative Ecological Niche Models for global change studies.

CENM are plagued of issues and uncertainties that have compromised their usefulness in global changes studies, especially regarding climate change (Wiens et al. 2009). Altogether, it has been claimed several times in the literature that CENM should be applied with caution, by carefully examining assumptions and uncertainties (Loiselle et al., 2003; Austin, 2007). For instance, Rocchini et al. (2011) suggested the generation of *maps of ignorance* as a *code of good practice* when developing CENM. After some overwhelming critics to CENM (see Cressey, 2008 on Nature news), one may think to what extent are these statistical models meaning anything at all, and to what extent they are useful. The studies involved in this work attempt to give some hints on whether information from these phenomenological models is acceptable, and to what extent we should rely on these models for global change projections.

For starters, it is especially meaningful to know the actual driving forces of species distributions, that is whether climate factors or others are constraining its current distribution and where climate may not be the reason for an absence of the species (Lobo et al., 2010). This is particularly relevant since most of studies about global change are directed to global warming. Ideally and in line with the theory, we would want to know how niche dimensions (B-A-M diagram) affect species and how they translate into space (and hopefully time; see Saupe et al. (2012) for implementation), in order to better understand constraints of species distribution. In study 2.1 we have shown a possible way to deal with this, by comparing the outputs from mechanistic ENM (eco-physiological models) and CENM probability scores at a specific regional scale. Essentially, we have linked the pattern (occurrence) and the process (growth) to robustly assess the degree of congruence between the two approaches. In this case, while the physiological model is linked to the potential niche of the species, CENM pictures the realized niche. Therefore the spatial differences between the two may

provide tests for further spatial hypothesis on species distributions (Buckley et al. 2010), as well as point out key factors that may or may not be exacerbated through global change dynamics. This is applicable to other kind of models aiming to determine the relative roles of climate and other factors across ranges (e.g. Angert 2009).

Ideally, if one is to model a change of state, an understanding of the process or mechanisms of change should be of paramount importance. In the realm of species distribution models (as a whole), it has been widely discussed the need for such understanding (Kearney, 2006). For instance, Kearney et al. (2008) successfully modeled cane toads in Australia without the need of any occurrence record, which exemplifies the power of understanding the underlying mechanisms in a modeling framework. I believe that the high dimensionality of processes involved in species distribution hampered the widespread use of some process-based models, in addition to the often limited information available to feed them. However, a mechanistic model does not need to be necessarily complex. For instance, thermal ecologists use some heat transfer functions that are easy to implement and they have predicted fairly well distributions of some organisms (Buckley et al. 2010) and patterns emerge between ectotherms and endotherms (Buckley et al., 2012). For instance, operative temperatures have been used in lizards or degree days for embryonic larval development in butterflies, or sunlight in plants. This illustrates the feasibility of potential comparisons between different approaches, without necessarily building very complex models. However, issues on the temporal resolution of input data (monthly averages versus daily) may largely affect its results (Kearney et al., 2012).

In global change studies, mechanisms are especially crucial for enhancing the reliability of projections of CENM. Modeling global change and in particular climate change, should not be detached from species physiology, especially in the case of plants. CENM in this case suffer from the lack of information on CO₂ concentration, simply because it is a global diffuse variable that it is assumed not to change spatially and therefore, it may not be a good estimator of the spatial distribution. However, in chapter 4 we have demonstrated that climate change is not only about precipitation and temperature, for which we recognize its crucial effects on vegetation, but it is also a global dynamism of biogeochemical cycles. In fact, our forecasts differed by the end of the century and revealed potential resilience of forests due to elevated CO₂ concentration. Using

development temperature of species from laboratory experiments, Buckley et al. (2011) published an interesting work in which they included this experimental parameterized variable in CENM. They concluded species-specific physiological traits may enhance predictions of climate change although their results were not conclusive. The power of including the mechanics is therefore evidenced and most importantly, whether included in CENM or compared to CENM's projections, they open up the avenue of biophysical dynamic mutual relationships of how species affect the environment and how the environment affect species.

In the present work I did not assess many other processes of utmost relevance that deserve special attention for geographical projection of species distribution like population dynamics, biotic interactions and human alterations. It has been already a while that population and meta-population processes including dispersal traits and biotic interactions have been incorporated to CENM, as pre-treatment data or post-treatment of results. This is particularly interesting for conservation studies, because species tracking climate change and persisting populations directly assess extinction or persistence, which are not necessarily included in the probability of presence and area outputs of from CENM (Fordham et al., 2012a).

Dispersal rates have been incorporated as a fixed term (Midgley et al., 2006), coupled with neighboring abundances (Iverson et al., 2004), but interestingly mechanistic models have also been developed for dispersal (Keith et al., 2008). This last case is especially meaningful because, dispersal rates are likely to change with changing conditions. However, in a recent work for European alpine plants (Dullinger et al., 2012), the incorporation of dispersal in CENM resulted in projections relatively similar to those assumed by the non-limited dispersal projections of CENM. This suggests that null vs. full dispersal assumptions may be a good strategy to bound dispersal scenarios. I believe these results may be very landscape dependent. The problem still relies in the fact that species presence data is also influenced by such mechanisms and therefore, decoupling the environmental (physiological) signal from others is not an easy task. In addition, Hui et al. (2012) warned us that indeed dispersal strategies are quite flexible and may vary regionally depending on the abiotic conditions, but in any case the addition of these dynamic approaches may improve predictions of range change (Keith et al., 2008; Zurell et al., 2009; Franklin, 2010).

Nevertheless, some of these processes and interactions have been, some would argue successfully, incorporated into CENM. Obviously, they may large changes in CENM projected climates (Bateman et al., 2012). Such is the case of biotic interactions, where the inclusion of the distribution of an interacting species may yield good results on performance and increase biological understanding (Heikkinen et al., 2007; Pellissier et al., 2010; Hof et al., 2012), or alternatively the use of a surrogate variable of biotic interactions (e.g. competition for light in Meier et al. 2010). Again, the issue relies on how to model the future distribution of a species depending on many other species and all of them depending on climate? How to disentangle the biotic from the abiotic signal? I think such practices may be very interesting when we want to determine to what extent biotic interaction constraint distributions, but may lead to some circularities in projections. On the other hand, they may point to some hints on patterns of co-occurrence and co-exclusion in the environmental space.

The avenue of biotic components inclusion could potentially lead to models that can accommodate such co-occurrence interactions (Meier et al., 2011), or determine the way in which a modifying feature of the environment (e.g. forest cover, fire, engineering species) modulates environmental variables (e.g. biotic modifiers; Linder et al., 2012). In any case, when sufficient data is available, there exist multiple ways by which it is potentially possible to incorporate multiple species interactions in CENM (e.g. error matrices in multivariate regression models (Kissling et al., 2011) or simultaneous equations). However, the enormous amount of biotic interactions (maybe most of them undiscovered) poses several challenges to the quantification of such effect. Surely, this avenue will represent a substantial step forward to the biological realism of these models (see Wisz et al. 2012 for a thoughtful review).

Another important process that I did not explore is human land use dynamics, despite accounting for a strong influence on biodiversity planning. A number of studies have addressed this question and (Bomhard et al., 2005; Eglington & Pearce-Higgins, 2012), in fact, most of land use and land cover variables may directly explain species distribution (Améztegui et al., 2010; Triviño et al., 2011) . I believe that the temporal scale of land use/ land change models and their projections (10 to 15 years maximum) precludes a direct comparison with those from climate change, which may operate at a coarser temporal scale. However, combined analysis provided valuable insights into

the synergies between these two effects (Lawson et al., 2010). Much less explored is the fact that actually these two components (land use and climate) interact and it may be difficult to disentangle the climatic changes due to land use shift or global warming (but see Clavero et al., 2011). To add more complexity we should not forget about disturbance events, where land use changes and expected extreme weather events interact and affect species distribution too (Regan et al., 2011; Conlisk et al., 2012).

Overall, many processes challenge the “snapshot” view of CENM. In a recent paper by Janet Franklin (2010), she calls for the development of dynamic approaches leaving behind some major limitations of modeling changes with, what she prefers to call, species distribution models (CENM in this work). In line with this, we have advanced in the dynamism by explicitly considering a measure of motion of the suitability index (chapter 7, bioclimatic velocity). Nevertheless, bioclimatic velocity is still constrained by the same issues of CENM. However, adding the temporal dimension yield valuable information to identify the temporal patterns of the rate of change between current and future situations. I feel that explicitly accounting for time is still an unresolved matter in CENM, although it may be crucial in decision-making and political processes. However, the question still remains whether CENM should account for transient states or on the contrary, we should leave this for explicitly dynamic models (see next section).

In the quest for finding general patterns or derive future global assessment of biodiversity and conservation networks, huge amounts of species have been modeled using CENM. For instance Garcia et al. (2012) published a study with more than 2000 species, depicting quite general patterns of change at continental scale (Africa). But can we model all species? It seems quite of a challenge taking into consideration that a lot of species are still to be discovered and there is a substantial lack of information on many species ecology (Linnean shortfall + Wallacean shortfall + Hutchinsonian shortfall³). However, attempts to scale up to higher levels of organizations (e.g. communities or ecosystems) are disputable since it is acknowledged that species respond individually to changes (McGeoch et al., 2006; Williams & Jackson, 2007) and spatial projections

³Linnean shortfall refers to the limitation that many species are yet to be discovered (Brown and Lomolino, 1998); Wallacean shortfall refers to the lack of information on species distribution (Lomolino, 2004); Hutchinsonian shortfall refers to the inadequate knowledge of the attributes that influence species niches.

may largely differ with individual versus aggregated approaches (chapter 5; but see dynamic community approaches in Mokany & Ferrier 2011). Essentially, perhaps the question is what biological level of organization may be suitable for capturing the environmental niche for a particular management goal. Empirical studies are reporting potential niche shifts in human dispersed species (e.g. invasive, Broennimann et al., 2007; Medley, 2010; da Mata et al., 2010), significant differences below the specific entity (Thompson et al., 2011) and new powerful in-depth data points out potential different responses of among populations of the same species (Benito Garzón et al., 2011). Moreover, other studies emphasize the differences in niche across species life stages (Quero et al., 2008).

I think that given the lack of information, modeling at the species level proved to be a good option, but if possible, the incorporation or ultimately the sole consideration of potential different responses below the specific level or different life stages, may be of major importance. Nevertheless, when the goal surpasses the species-specific levels, using broad climatic classifications or higher-taxon levels are also useful. For instance, (Thuiller et al., 2005) measured the potential of the South African flora (in general) to become invasive in different regions. The aim in that case was to derive general invasive risk and therefore such approach may be appropriate.

In any case, data is a crucial ingredient of our models. For instance, some global data may not achieve the desirable spatial resolution for regional studies (Bedia et al., in press), however some attempts have been performed to downscale coarse data to a desirable working scale (e.g. downscaling atlas data (Bombi & D'Amen, 2012; Niamir et al., 2011)). Whether downscaled or not, we need to have the maximum amount of data, at least before a quality control phase. In a very illustrative paper, Lobo (2008) examined the effects of data versus modeling techniques. The results of the study pointed that we may need to focus more on biologically meaningful data than in high-profile modeling techniques. Accordingly, new calls for global integration of species distribution data are emerging (Jetz et al., 2012).

In any case, data quality is essential to apply CENM. Usually, quality data is available only at local or regional domains but luckily the emergence of geoportals through the World Wide Web has enabled the organization of contents and services (geosearch tools, community information,

support resources, data and applications). Both geoportals and spatial infrastructure data (SID) have simplified its acces to the geographic information, enabling that we can have relatively high quality information for extensive geographic coverages (e.g. continental level or even world). Such is the case of initiatives such as GBIF (Global Biodiversity Information Facility, <http://www.gbif.org>), WORLDCLIM (<http://www.worldclim.org/>), the Global Index of Vegetation-Plot Databases (<http://www.givd.info>) or the Global Observation Research Initiative in Alpine environments (<http://www.gloria.ac.at>), to mention a few.

However, we are still facing interoperability and quality challenges. Most metadata are still lacking of quality parameters and the disparity of data formats may produce an onerous work of data gathering. Therefore it may be difficult to address the appropriateness of a given GIS layer to be used in a CENM. To address this problem, initiatives as the Open Geospatial Consortium are collaborating to develop standards for distributed geospatial computing, together with efforts with those of data portal builders (e.g. GBIF).

This lack of quality data is especially interesting in a moment where a lot of methodological refinements are sprouting out (e.g. new techniques, data selection procedures, autocorrelation control methods, etc.). However, little funding is actually dedicated to improve our knowledge on biodiversity, which could lead to biased results of nature conservation design to those over-monitored areas (Ahrends et al., 2011). In addition, less scientific reward is given on those data producers (Chavan & Penev, 2011) and consequently, it is not surprising that less people are actually trained to survey, recognize and classify species (Wägele et al., 2011). Despite the need of such methodological improvements, I think the role of data acquisition of biodiversity data has been downturned and needs to be reemphasized and mostly, temporarily connected to modeling efforts. Ideally, it would be desirable to set up a dynamically linked process of data acquisition and CENM resulting maps, even enabling the calibration of a mechanistic component (De Cáceres & Brotons, 2012). A constant feedback is then set, so that maps inform potential distribution and that enables new campaigns of data acquisition. Volunteering and undergraduate students may be also increase distribution databases and to recover the role of taxonomy (Pearson et al., 2011). For instance, geositional applications in cellular devices (e.g. ZamiaDroid, <http://biodiver.bio.ub.es/>) may indeed open up as a new opportunity to acquire new information or validate existing geodata

bases (e.g. geowiki). However, expert knowledge and metadata handling will have to be undertaken in order to carefully assess the quality of such new databases, as potential taxonomic errors may occur. Although barcoding may be a good technique to overcome such errors, we are still at an early-stage for its widespread use as several technical issues still remain unresolved, especially in plants (Chase & Fay, 2009).

In addition to biological data, some crucial environmental data are still lacking. For instance, soil types and soil properties are not easily available in some countries (e.g. Spain) although this may greatly affect our predictions of species distributions, especially plants whose relationship with soil is really tight. Sometimes, soil data exist at a very coarse resolution (e.g. Joint Research Centre, <http://eusoils.jrc.ec.europa.eu/>), or at high resolution but low coverage (e.g. only for agricultural areas). We acknowledge the difficulty of building such information but happily great efforts are being made (e.g. Catalonia Soil Map). Further, remote sensing missions like SMOS (Soil Moisture Ocean Salinity; European Space Agency) may improve our understanding and descriptions of the spatial distributions of soils and its properties.

After all, *are CENM worth for global change studies despite their known caveats?* I think so. Let's first recall a developed vulnerability framework for species coping with climate change, which is one of the particular cases in which CENM have been widely applied. Dawson et al. (2011) proposed a vulnerability framework for climate change impacts on species distributions, in which vulnerability encompasses three features: exposure (extent of climate change likely to be experienced by a species), sensitivity (degree of dependency between current climate and species/population persistence or performance) and adaptive capacity (ability to cope with new environments in situ or migrate to other suitable environments). Adaptive capacity and sensitivity may be difficult to address for many species since they rely on detailed knowledge of genetic structure, demography, species physiology or life history, for which data are unfortunately largely lacking. However, exposure to climate change may be estimated with species distribution data by using CENM on the target species and is useful for bounding the range of actions decision makers may need to consider.

It is also important to stress here that other authors have developed these three features of vulnerability (exposure, adaptive capacity and sensitivity after Dawson et al. 2011) using only CENM (Summers et al., 2012; Crossman et al., 2011). For instance, they used the sum of the absolute value of probabilities of CENM in current and projected distribution to derive species sensitivity weights. The same authors have referred to adaptive capacity assessing only cost-dispersal routes. Although I appreciate the full potentiality of this approach, I think it is a very restrictive use of the vulnerability framework. First, CENM correlative approach is not designed to account for adaptive capacity, which is dependent not on the phenomenon of occurrence, but from species traits itself. Second, sensitivity and exposure are measures widely overlapped under this approach. I think that adaptive capacity and sensitivity would better benefit from experimental or mechanistic approaches, or at least in combination with CENM.

In summary, CENM assess the degree of change from known occurrences to current or new environmental conditions. However, the degree to which this probability of occurrence informs a certain process of interest (e.g. invasion, migration, extinction, etc.) will largely rely on data treatment and availability, statistical technique and the possibility of contrasting with alternative modeling approaches. The reader may at this point feel disappointed, but as noted during this work, picturing similar conditions to known frequency of species occurrences has a tremendous wide range of applications and constitutes a powerful hypothesis generator.

Indeed it is likely that a researcher or a manager feel overwhelmed amidst the complexities of biological processes that need to be considered in the face of global change and even only considering climate change (Bonan, 2008; Heimann & Reichstein, 2008). On one hand, the scientific community urges managers and policy makers to develop conservation plans and policies that are in line with projected climate and land use changes through the end of the century. On the other hand, scientists have demonstrated little success in transforming predictions (especially climate change predictions) into policy-ready information with an acceptable range of uncertainty (Kerr, 2011). For instance, assessments of vulnerability for a broad range of species that consider multiple processes are still lacking due to a paucity of detailed information for many taxa, even though projections of threats are fundamental to conservation management plans. I do not underestimate here large collaborative networks of scientists that are developing useful tools, but it is

acknowledged that we are still at early stage of understanding, especially understanding each other (scientist and decision makers).

Under these circumstances, *are CENM for global change informative for practitioners?* I guess we have to find the answer in those who have been applying CENM for long in the management scene. Iverson and colleagues, from the United States Forest Service, have recently published a very thoughtful work entitled: *“lessons learned while integrating habitat, dispersal, disturbance and life-history traits into species habitat models under climate change”* (Iverson et al., 2011). In that paper, summarizing common practices and new approaches, they developed a set of modification factors framework (ModFacs; see also Matthews et al., 2011) based on expert knowledge, together with other modeling efforts (dynamic dispersal module: SHIFT). It is very interesting in this case that expert knowledge, when wisely handled, may modify projections from CENM in a useful way for managers. However they also recommend to *“Use species models as guidelines for regional trends. Because of uncertainties and scale, they are not usually appropriate for fine-scale management without the regional context”*.

8.2 The ongoing and future modeling scene and the role for CENM

CENM are still useful and will likely to continue being an essential tool for natural scientists dealing with species distributional changes, particularly because a lot of biological data is still lacking to develop alternative modeling tools. Ongoing studies outline better strategies dealing with potential issues of data and new CENM algorithms may show up in the forthcoming years. Overall, better practices in the application of CENM are being implemented nowadays (Anderson & Gonzalez Jr., 2011).

Meanwhile, we are on the way of understanding and assessing the strengths and weaknesses between modeling approaches (e.g. correlative versus process-based). Surprising as it may seem, to my knowledge the existing comparisons have especially addressed the issue of climate change (Hijmans & Graham, 2006; Buckley, 2008; Morin & Thuiller, 2009; Buckley et al., 2010; Keenan et

al., 2011). In a recent paper, Cheaib et al. (2012) performed a very complete assessment of uncertainties using 8 models ranging from correlative to process-based models, to predict tree range changes in France due to climate change. The authors evidenced regions of high uncertainty and point out in which way taking into account a certain process affect its outcomes.

In my opinion, there is not much of a point in praising the strengths of one approach over another, but rather make them work in tandem. We need to elucidate potential synergies in an open-minded modeling framework. Thereby, I feel that the ongoing modeling scene has shifted towards a hybrid modeling of species distribution (sensu Dormann et al. 2012). Hybrid models may be defined as the “sequential application of different models” and they may represent a useful way in order to “capitalize on the strength and advantage of both approaches and concepts to make more reliable and useful predictions” (Gallien et al., 2010).

But what is the role of CENM under this modeling framework? Happily, readily available examples can be found in the literature. In general, coupling CENM suitability with other models has been realized in the context of dispersal and migration process, by balancing the forecasts of future potential habitats (Thuiller et al., 2008). In some cases a cellular automaton approach is built where habitat suitability represent environmental conditions of a state. For instance, Wilson et al. (2009) used habitat suitability as a surrogate for extinction in their coupled dynamic model of migration. In many cases, probability of occurrence is assumed to be a surrogate of carrying capacity or intrinsic growth rate (Pagel & Schurr, 2011). Anderson et al., (2009) used climatic suitability maps to derive carrying capacity at each time step in their metapopulation models by taking into account suitable patches and the area of those patches (after thresholding).

In general, CENM have been used to provide potential suitable habitat in a given time-step, in which to embed a certain process. They provide the spatial structure of patches of potential habitats (Regan et al., 2011; Fordham et al., 2012b; Conlisk et al., 2012) and, in simulation experiments they may be combined with different dynamic processes (e.g. disturbance, etc.), suitable habitat patches are the ones susceptible of being colonized (Iverson et al., 2004; Engler & Guisan, 2009). Another fruitful example of this is the BioMove framework (Midgley et al., 2010), an integrated modeling platform that simulates several processes (e.g. disturbance, dispersal,

succession, etc.) at the plant functional type level. In this case, suitability is used (optionally) not only to exclude some processes (e.g. setting off recruitment in unsuitable habitats) but also to scale them.

To sum up, we pointed out several manners by which to embed CENM results in a broader modeling framework. However, I feel that it may be risky to use suitability index as a surrogate of population or individual dynamic performance traits. Some authors have evidenced a weak (or null) link between suitability and abundance (Sagarin et al., 2006) or suitable area as a surrogate of extinction risk (Zurell et al., 2012). Further, as pointed out by Zurell (2009), it is often assumed that suitability index represents resources, shelter or at least the available climate space. However, it is difficult to know the full potential of the suitability index in order to be linked with several other processes, especially because we do not know what low values of suitability mean (unsuitability areas). I envision future refinements in models to explicitly state what the suitability or probability of occurrence is modeling in each geographic region of interest, so we can derive better hybrid models (Saupe et al., 2012). In other words, a spatially explicit biological refinement of the probability of presence.

Basically, these hybrid approaches rely on a framework of spatially nested models. A good example is the recent work by Boulangeat et al. (2012) in which they applied a theoretical framework top-down approach, of presence-absence-model and an abundance model. In this work they fitted many processes (e.g. dispersal, biotic interactions, etc.) and managed to both explain occurrence and abundance. As in many other works of new hybrid models, CENM are aimed to be the “physiological filter”, the potential of the abiotic factors. I concur with this sequential approach, but I would rather emphasize here that we should perform a post-treatment of suitability, maybe by comparing it to simple process-based models, in order to refine the physiological signal of suitability outputs.

However, it is not only suitability that should be accurately examined. In fact, in recent analysis of uncertainty in these sort of hybrid models, Zurell et al. (2012) pointed out that larger sources of uncertainty were concentrated rather in dynamic components than in climate change scenario or CENM algorithm. Furthermore, it is important to highlight that the calibration of the mechanistic

component may also be highly influenced by its calibration procedure and the data used for that aim (De Cáceres & Brotons, 2012).

At broader scales and similarly to such sequential approaches we can find Dynamic Global Vegetation Models (DGVM). These process-based models represent a scaling-up approach of processes starting from physiology until ecosystem dynamics. First developments treated vegetation as a green cover, but they are currently using plant functional types to address vegetation heterogeneity within each tile of analysis. These (meta-) models could be classified as either fitted process-based models or forward process-based models. They offer the potential to determine broad patterns of future vegetation types and transient responses to global change including biogeochemical cycles. However, uncertainty analysis and treatment of tile heterogeneity are still largely lacking. I envision potential use of CENM to inform DGVM in at higher spatial resolution in order to determine future species distributions, or at least hypothesize future regional contexts within each tile (I actually found a new study while proof reading this phd dissertation: Notaro et al., 2012).

Finally, whether nested hybrid models or fitted-process based models, such nested meta-models will have to explore potential connections of scale and how bottom up and top-down controls inform different mechanisms, and most intriguing is to identify at which temporal resolution.

8.3 So far, what have I learned along the way? Bits and pieces about CENM in a nutshell.

“ The most common ‘error’ of any modeller is to ‘believe’ a model “

(Dormann et al., 2012)

Do not be afraid of comparing apples and oranges!

Model inter-comparison greatly facilitates the understanding of species distributions and improves the assessment of global change projections. We should enhance the mix and comparison between modeling approaches, as they may shed light of potential caveats while open up new avenues of improvement.

Fit the model for the purpose and, if possible, do not run out of time.

The bioclimatic profiling (niches) is dynamic and the identification is contingent to the biological level of organization. For global change studies, CENM need to carefully consider their working spatial and biological scale in response to a desired management or scientific framework. Adding the temporal scale to predictions may better picture spatiotemporal dynamics, although they will be constrained by the static nature of their responses.

Watch out going beyond your predictions

Great care should be taken when considering the probability of presence as a surrogate of physiological or population performance. For global change studies, CENM measure exposure to new conditions. In some case, it may be a surrogate of other biological factors, but not necessarily. As a consequence, changes in the bioclimatic profiling should be interpreted as exposure to new conditions unless the biological meaning of probability of presence has been assessed.

8.4 References

- Ahrends A, Burgess ND, Gereau RE et al. (2011) Funding begets biodiversity. *Diversity and Distributions*, **17**, 191-200.
- Améztegui A, Brotons L, Coll L (2010) Land-use changes as major drivers of mountain pine (*Pinus uncinata* Ram.) expansion in the Pyrenees. *Global Ecology and Biogeography*, **19**, 632-641.
- Anderson BJ, Akçakaya HR, Araújo MB, Fordham DA, Martinez-Meyer E, Thuiller W, Brook BW (2009) Dynamics of range margins for metapopulations under climate change. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1415-1420.
- Anderson RP, Gonzalez Jr. I (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, **222**, 2796-2811.
- Angert AL (2009) The niche, limits to species' distributions, and spatiotemporal variation in demography across the elevation ranges of two monkeyflowers. *Proceedings of the National Academy of Sciences*, **106**, 19693-19698.
- Austin M (2007) Species distribution models and ecological theory: A critical assessment and some possible new approaches. *Ecological Modelling*, **200**, 1-19.
- Bateman BL, VanDerWal J, Williams SE, Johnson CN (2012) Biotic interactions influence the projected distribution of a specialist mammal under climate change. *Diversity and Distributions*, in press.
- Bedia J, Herrera S, Gutierrez JM (in press) Climate-change projection of species distributions: Sensitivity to baseline climatology and regional. *Global and Planetary Change*, in press.
- Benito Garzón M, Alía R, Robson TM, Zavala MA (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766-778.
- Bombi P, D'Amen M (2012) Scaling down distribution maps from atlas data: a test of different approaches with virtual species. *Journal of Biogeography*, **39**, 640-651.

- Bomhard B, Richardson DM, Donaldson JS et al. (2005) Potential impacts of future land use and climate change on the Red List status of the Proteaceae in the Cape Floristic Region, South Africa. *Global Change Biology*, **11**, 1452-1468.
- Bonan GB (2008) Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science*, **320**, 1444-1449.
- Boulangeat I, Gravel D, Thuiller W (2012) Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters*, **15**, 584-593.
- Broennimann O, Treier UA, Müller-Schärer H, Thuiller W, Peterson A, Guisan A (2007) Evidence of climatic niche shift during biological invasion. *Ecology Letters*, **10**, 701-709.
- Buckley LB (2008) Linking traits to energetics and population dynamics to predict lizard ranges in changing environments. *The American Naturalist*, **171**, E1-E19.
- Buckley LB, Waaser SA, MacLean HJ, Fox R (2011) Does including physiology improve species distribution model predictions of responses to recent climate change? *Ecology*, **92**, 2214-2221.
- Buckley LB, Urban MC, Angilletta MJ, Crozier LG, Rissler LJ, Sears MW (2010) Can mechanism inform species? distribution models? *Ecology Letters*, **13**, 1041-1054.
- Buckley LB, Hurlbert AH, Jetz W (2012) Broad-scale ecological implications of ectothermy and endothermy in changing environments. *Global Ecology and Biogeography*, in press.
- Chase MW, Fay MF (2009) Barcoding of plants and fungi. *Science*, **325**, 682-683.
- Chavan V, Penev L (2011) The data paper: a mechanism to incentivize data publishing in biodiversity science. *BMC Bioinformatics*, **12**, S2.
- Cheaib A, Badeau V, Boe J et al. (2012) Climate change impacts on tree ranges: model intercomparison facilitates understanding and quantification of uncertainty. *Ecology Letters*, **15**, 533-544.
- Clavero M, Villero D, Brotons L (2011) Climate change or land use dynamics: do we know what climate change indicators indicate? *PLoS One*, **6**, e18581.
- Conlisk E, Lawson D, Syphard AD, Franklin J, Flint L, Flint A, Regan HM (2012) The Roles of Dispersal, Fecundity, and Predation in the Population Persistence of an Oak (*Quercus engelmannii*) under Global Change. *PLoS ONE*, **7**, e36391.

- Crossman ND, Bryan BA, Summers DM (2011) Identifying priority areas for reducing species vulnerability to climate change. *Diversity and Distributions*, **18**, 60-72.
- Cressey D (2008) Pushing the modeling envelope. Nature news. 15th September 2008. doi:10.1038/news.2008.1108
- da Mata RA, Tidon R, Côrtes LG, De Marco P, Diniz-Filho JAF (2010) Invasive and flexible: niche shift in the drosophilid *Zaprionus indianus* (Insecta, Diptera). *Biological Invasions*, **12**, 1231-1241.
- Dawson TP, Jackson ST, House JI, Prentice IC, Mace GM (2011) Beyond Predictions: Biodiversity Conservation in a Changing Climate. *Science*, **332**, 53-58.
- De Cáceres M, Brotons L (2012) Calibration of hybrid species distribution models: the value of general-purpose vs. targeted monitoring data. *Diversity and Distributions*, in press.
- Dormann CF, Schymanski SJ, Cabral J et al. (2012) Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography*, in press.
- Dullinger S, Gatttringer A, Thuiller W et al. (2012) Extinction debt of high-mountain plants under twenty-first-century climate change. *Nature Climate Change*, in press.
- Eglinton SM, Pearce-Higgins JW (2012) Disentangling the Relative Importance of Changes in Climate and Land-Use Intensity in Driving Recent Bird Population Trends. *PloS one*, **7**, e30407.
- Engler R, Guisan A (2009) MigClim: predicting plant distribution and dispersal in a changing climate. *Diversity and Distributions*, **15**, 590-601.
- Fordham DA, Resit Akçakaya H, Araújo MB et al. (2012a) Plant extinction risk under climate change: are forecast range shifts alone a good indicator of species vulnerability to global warming? *Global Change Biology*, **18**, 1357-1371.
- Fordham DA, Watts MJ, Delean S, Brook BW, Heard LMB, Bull CM (2012b) Managed relocation as an adaptation strategy for mitigating climate change threats to the persistence of an endangered lizard. *Global Change Biology*, in press.
- Franklin J (2010) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, **16**, 321-330.

- Gallien L, Münkemüller T, Albert CH, Boulangeat I, Thuiller W (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, **16**, 331-342.
- Garcia RA, Burgess ND, Cabeza M, Rahbek C, Araújo MB (2012) Exploring consensus in 21st century projections of climatically suitable areas for African vertebrates. *Global Change Biology*, **18**, 1253-1269.
- Heikkinen RK, Luoto M, Virkkala R, Pearson RG, Körber JH (2007) Biotic interactions improve prediction of boreal bird distributions at macrocales. *Global Ecology and Biogeography*, **16**, 754-763.
- Heimann M, Reichstein M (2008) Terrestrial ecosystem carbon dynamics and climate feedbacks. *Nature*, **451**, 289-292.
- Hijmans RJ, Graham CH (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272-2281.
- Hof AR, Jansson R, Nilsson C (2012) How biotic interactions may alter future predictions of species distributions: future threats to the persistence of the arctic fox in Fennoscandia. *Diversity and Distributions*, **18**, 554-562.
- Hui C, Roura-Pascual N, Brotons L, Robinson RA, Evans KL (2012) Flexible dispersal strategies in native and non-native ranges: environmental quality and the “good-stay, bad-disperse” rule. *Ecography*, in press.
- Iverson LR, Prasad AM, Matthews SN, Peters MP (2011) Lessons learned while integrating habitat, dispersal, disturbance, and life-history traits into species habitat models under climate change. *Ecosystems*, **14**, 1005-1020.
- Iverson LR, Schwartz MW, Prasad AM (2004) How fast and far might tree species migrate in the eastern United States due to climate change? *Global Ecology and Biogeography*, **13**, 209-219.
- Jetz W, McPherson JM, Guralnick RP (2012) Integrating biodiversity distribution knowledge: toward a global map of life. *Trends in Ecology & Evolution*, **27**, 151-159.
- Kearney MR, Matzelle A, Helmuth B (2012) Biomechanics meets the ecological niche: the importance of temporal data resolution. *The Journal of experimental biology*, **215**, 922-933.

- Kearney M (2006) Habitat, environment and niche: what are we modelling? *Oikos*, **115**, 186-191.
- Kearney M, Phillips BL, Tracy CR, Christian KA, Betts G, Porter WP (2008) Modelling species distributions without using species distributions: the cane toad in Australia under current and future climates. *Ecography*, **31**, 423-434.
- Keenan T, Serra JM, Lloret F, Ninyerola M, Sabaté S (2011) Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO2 matters! *Global Change Biology*, **17**, 565-579.
- Keith DA, Akçakaya HR, Thuiller W et al. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560-563.
- Kerr RA (2011) Time to Adapt to a Warming World, But Where's the Science? *Science*, **334**, 1052-1053.
- Kissling WD, Dormann CF, Groeneveld J et al. (2011) Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, in press.
- Lawson DM, Regan HM, Zedler PH, Franklin J (2010) Cumulative effects of land use, altered fire regime and climate change on persistence of *Ceanothus verrucosus*, a rare, fire-dependent plant species. *Global Change Biology*, **16**, 2518-2529.
- Linder P, Bykova O, Dyke J et al. (2012) Biotic modifiers, environmental modulation and species distribution models. *Journal of Biogeography*, in press .
- Lobo JM (2008) More complex distribution models or more representative data? *Biodiversity Informatics*, **5**, 14-19.
- Lobo JM, Jiménez-Valverde A, Hortal J (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103-114.
- Loiselle BA, Howell CA, Graham CH, Goerck JM, Brooks T, Smith KG, Williams PH (2003) Avoiding Pitfalls of Using Species Distribution Models in Conservation Planning. *Conservation Biology*, **17**, 1591-1600.

- Matthews SN, Iverson LR, Prasad AM, Peters MP, Rodewald PG (2011) Modifying climate change habitat models using tree species-specific assessments of model uncertainty and life history-factors. *Forest Ecology and Management*, **262**, 1460-1472.
- McGeoch MA, Le Roux PC, Hugo EA, Chown SL (2006) Species and community responses to short-term climate manipulation: Microarthropods in the sub-Antarctic. *Austral Ecology*, **31**, 719-731.
- Medley KA (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.
- Meier ES, Edwards Jr TC, Kienast F, Dobbertin M, Zimmermann NE (2011) Co-occurrence patterns of trees along macro-climatic gradients and their potential influence on the present and future distribution of *Fagus sylvatica* L. *Journal of Biogeography*, **38**, 371-382.
- Meier ES, Kienast F, Pearman PB, Svenning J, Thuiller W, Araújo MB, Guisan A, Zimmermann NE (2010) Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, **33**, 1038-1048.
- Midgley GF, Hughes GO, Thuiller W, Rebelo AG (2006) Migration rate limitations on climate change-induced range shifts in Cape Proteaceae. *Diversity and Distributions*, **12**, 555-562.
- Midgley GF, Davies ID, Albert CH et al. (2010) BioMove ? an integrated platform simulating the dynamic response of species to environmental change. *Ecography*, **33**, 612-616.
- Mokany K, Ferrier S (2011) Predicting impacts of climate change on biodiversity: a role for semi-mechanistic community-level modelling. *Diversity and Distributions*, **17**, 374-380.
- 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.
- Niamir A, Skidmore AK, Toxopeus AG, Muñoz AR, Real R (2011) Finessing atlas data for species distribution models. *Diversity and Distributions*, **17**, 1173-1185.
- Notaro M, Mauss A, Williams JW (2012) Projected vegetation changes for the American Southwest: combined dynamic modeling and bioclimatic-envelope approach. *Ecological Applications*, **22**, 1365-1388.
- Pagel J, Schurr FM (2011) Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography*, **21**, 293-304.

- Pearson DL, Hamilton AL, Erwin TL (2011) Recovery plan for the endangered taxonomy profession. *Bioscience*, **61**, 58-63.
- Pellissier L, Anne Bråthen K, Pottier J et al. (2010) Species distribution models reveal apparent competitive and facilitative effects of a dominant species on the distribution of tundra plants. *Ecography*, **33**, 1004-1014.
- Quero JL, Gómez-Aparicio L, Zamora R, Maestre FT (2008) Shifts in the regeneration niche of an endangered tree (*Acer opalus* ssp. *granatense*) during ontogeny: Using an ecological concept for application. *Basic and Applied Ecology*, **9**, 635-644.
- Regan HM, Syphard AD, Franklin J, Swab RM, Markovchick L, Flint AL, Flint LE, Zedler PH (2011) Evaluation of assisted colonization strategies under global change for a rare, fire-dependent plant. *Global Change Biology*, **18**, 936-947.
- Rocchini D, Hortal J, Lengyel S, Lobo JM, Jiménez-Valverde A, Ricotta C, Bacaro G, Chiarucci A (2011) Accounting for uncertainty when mapping species distributions: the need for maps of ignorance. *Progress in Physical Geography*, **35**, 211-226.
- Sagarin RD, Gaines SD, Gaylord B (2006) Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology & Evolution*, **21**, 524-530.
- Saupe EE, Barve V, Myers CE et al. (2012) Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling*, **237–238**, 11-22.
- Summers DM, Bryan BA, Crossman ND, Meyer WS (2012) Species vulnerability to climate change: impacts on spatial conservation priorities and species representation. *Global Change Biology*, **18**, 2335-2348.
- Thompson GD, Robertson MP, Webber BL, Richardson DM, Le Roux JJ, Wilson JR (2011) Predicting the subspecific identity of invasive species using distribution models: *Acacia saligna* as an example. *Diversity and Distributions*, **17**, 1001-1014.
- Thuiller W, Albert C, Araújo MB et al. (2008) Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 137-152.

- Thiller W, Richardson DM, Pyzek P, Midgley GF, Hughes GO, 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.
- Triviño M, Thuiller W, Cabeza M, Hickler T, Araújo MB (2011) The Contribution of Vegetation and Landscape Configuration for Predicting Environmental Change Impacts on Iberian Birds. *PloS one*, **6**, e29373.
- Wägele H, Klusmann-Kolb A, Kuhlmann M, Haszprunar G, Lindberg D, Koch A, Wägele JW (2011) The taxonomist-an endangered race. A practical proposal for its survival. *Frontiers in Zoology*, **8**, 25.
- Wiens JA, Stralberg D, Jongsomjit D, Howell CA, Snyder MA (2009) Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, **106**, 19729-19736.
- Williams JW, Jackson ST (2007) Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, **5**, 475-482.
- Wilson RJ, Davies ZG, Thomas CD (2009) Modelling the effect of habitat fragmentation on range expansion in a butterfly. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 1421-1427.
- Wisz MS, Pottier J, Kissling WD et al. (2012) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, in press.
- Zurell D, Grimm V, Rossmannith E, Zbinden N, Zimmermann NE, Schröder B (2012) Uncertainty in predictions of range dynamics: black grouse climbing the Swiss Alps. *Ecography*, **35**, 590-603.
- Zurell D, Jeltsch F, Dormann CF, Schröder B (2009) Static species distribution models in dynamically changing systems: how good can predictions really be? *Ecography*, **32**, 733-744.

AGRAÏMENTS I RECONeixEMENTS

Agraeixo a tota la meva família haver-me recolzat física i moralment en aquest esgotador viatge (sí comptem la carrera, també recolzament monetari). Indubtablement això també va dedicat al Murjul, que ha estat un company incansable i un suport incomensurable especialment aquest darrer tram, gràcies pels dibuixos, els sopars i la paciència. Dono les gràcies a en Miquel, que em va iniciar i em va rescatar del món privat per a emprendre aquest treball d'aprenentatge constant que destijo poder continuar. Així, no m'oblido dels companys i companyes de Botànica, que m'han ajudat i alhora m'han picat per poder fer una feina millor. Impossible oblidar aquelles primeres classes d'identificació de plantes...sense l'ajuda de l'Ainhoa, l'Alex i l'Encarni no me n'hagués sortit (això és per a vosaltres també). Gent de pol-len, això també us ho dec a vosaltres, gràcies pels riures (i crits, ejem...) de la cuina de 2n pis, fa que el passadís sigui més interessants. Conchiiii!gracias por la inovación docente y esas actividades inventadas!Al mandala, per haver-me fet explicar 200mil vegades el que estava fent, i no entendre-ho!Als raterinyos, més presents al principi d'aquesta tesi però no per això menys importants... sempre redcordaré aquells migdies (i tardes, ejem...) a la gespa..també això és part del procés d'aprenentatge. Viri!neeeen!q al final siré doctor i tot!jajajja, gràcies per tantes festes i tants esmorzars plegats,..fa que la uab sigui també com una espècie de casa familiar. Trevlthanks for your friendship, support and wise comments, you made me grow as a person & scientist...although I've never made it to play the djidji ☺! Mariaaa!merci per tot (tb la portada) però per haver viscut moltes coses plegats i fer-m'ho pasar tant bé als congressos..siempre no quedará Iruña. Lasse...al final nada del "love ecology"...otra vez será pero gracias por estrujarme al cerebro con ideas raras y con R. Rebeca, merci pel teu superpositivisme... el cafè sap millor quan et passes per botànica! Leiretxu, poco en la uab pero mucho fuera ... gracias por todo lo cerca que has estado, especialmente en Geografía. Guille, merci per alguna que altra xerrada científica, ni que fos masturbació mental, ha estat molt plaent, jjajajaja. Gent del CREAM (no dic noms per por d'oblidar-me algú, però que tothom es doni per al-ludit) merci per amenitzar, dinamitzar i provocar tants sentiments entre al gent que estem fent el doctorat..o el que sigui. Biojuligans sou geniaaals!no sé si teniu alguna relació amb la tesi però us envio una abraçada des d'aquest moment d'enajenació mental. Alex! .bat-killer...va ser un plaer seure al teu costat i aprendre tant...ets un crack!Txeeee!!!!gRRRRRràcies per tot!m'he sentit menys fRiki al teu costat i m'has ajudat molt amb la Cesca, entre d'altres! Grampaires de tots colors i paisatges!he après coses més importants que les acadèmiques amb vosaltres. Albert (sí sí, tu, el Pèlachs) gràcies per fer-me describir tantes altres cares de les muntanyes. Paco, sense tu aquesta recerca no hagués pres el ritme que ha pres, gàcies pels teus savis comentaris i el teu suport. CrisCeaaa, a part de núvols,..crec q ara vaig tant ràpid amb el teclat gràcies a les teves lliçons! Jordi C. Ets un crack, moltes gràcies per grans consells professionals i per fer-me practicar el processament d'explicacions a 200km/h. Jorgitooooo!AZ no hubiera sido lo mismo sin ti, pinche mexa...eres genial y eso no se resume aquí, hiciste q mi experiencia ahi fuera lo mejor posible. Marceeeeee! :) acaba tu tesis ya este momento loco de agradecimiento vale la pena, no sabés lo q me hace sonreir acordarme de ti..esta tesis también se acaba mejor gracias a ti. Lau&Lara o Lau&Lara, gracias por amenizar tantos encuentros en Tempe, .. y por aprender q me puede gustar el Fernet (Fernes?). Bray! A ti te debo mucho del apoyo en el Lab ... y por ser de las pocas personas que entendía de lo que estaba hablando! Jullie, thank you for being there by my side and making the lab much more joyful!Janet, you were an incredible supervisor there, thank you for making me grow as a scientist. Tronk@s de Madriz gracias por vuestras múltiples recibimientos, ahí y en Évora! Maria (Triviño) ha sido genial verte durante este ultimo año, gracias por tu acogida! Edu, it was fun sharing my Évora experience with you..I will always remember that BeachLabMeeting and some projects yet to be terminated ;) ! Raquel,muito obrigado por hacerme sentir agusto en donde sea q estas. Miguel, muito obrigado por hacerme estrujar el cerbro y enseñarme que leer mucho es la base para decir cosas interesantes. André, jij dachte dat ik konde niet maar uiteindelijk het is the end my friend! Mieke, bedankt voor de nachts and leuk conversaties in BCN and making me think I was doing the right thing with this phd thesis. César (o Sésar), Lio y Rudi...estos últimos meses en la biblio han sido muy divertidos gracias a vosotros. En darrer terme, vull agrair a la gran quantitat de persones que m'han ajudat desinteressadament a través de les llistes R-SIG-GEO i Biomod (especialment en Wilfred), sense elles això no hagués estat possible. També vull reconeixer l'esforç que tanta i tanta gent està fent per a desenvolupar softwares lliures del que aquesta recerca n'ha tret partit. Així mateix, no puc acabar sense reconeixer la UAB, que amb la seva política de beques d'investigació ha fet que això arribi a bon port.

APPENDIX

APPENDIX 1

APPENDIX 1.1. Literature extension

The present literature extension orders relevant literature reviewed during this PhD thesis. It is not the intention here to perform a complete literature review here but rather to illustrate the manuscript that have inspired this research. The reader may find the classification proposed to be overlapping in several aspects. We acknowledge the caveats of classifying manuscripts, especially because most of them may deal with different research questions. However, we argue this classification is valuable and certainly organizes the large amount of information that is being produced in the fast-moving species distribution modelling discipline.

SPECIES DISTRIBUTION RECORDS

Foody GM (2011) Impacts of imperfect reference data on the apparent accuracy of species presence-absence models and their predictions. *Global Ecology and Biogeography*, **20**, 498-508.

Graham CH, Elith J, Hijmans RJ, Guisan A, Townsend Peterson A, Loiselle BA, The Nceas Predicting Species Distributions Working Group (2008) The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, **45**, 239-247.

Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S, Peterson AT (2007) What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs*, **77**, 615-630.

Hortal J, Lobo JM, Jiménez-Valverde A (2007) Limitations of Biodiversity Databases: Case Study on Seed-Plant Diversity in Tenerife, Canary Islands. *Conservation Biology*, **21**, 853-863.

Kent R, Carmel Y (2011) Presence-only versus presence-absence data in species composition determinant analyses. *Diversity and Distributions*, **17**, 474-479.

Li W, Guo Q, Elkan C (2011) Can we model the probability of presence of species without absence data? *Ecography*, no-no.

Geositional errors and scaling

Bombi P, D'Amen M (2012) Scaling down distribution maps from atlas data: a test of different approaches with virtual species. *Journal of Biogeography*, **39**, 640-651.

Gastón A, García-Viñas JI (2010) Updating coarse-scale species distribution models using small fine-scale samples. *Ecological Modelling*, **221**, 2576-2581.

Marcer A, Pino J, Pons X, Brotons L (2012) Modelling invasive alien species distributions from digital biodiversity atlases. Model upscaling as a means of reconciling data at different scales. *Diversity and Distributions*, n/a-n/a.

Naimi B, Skidmore AK, Groen TA, Hamm NAS (2011) Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. *Journal of Biogeography*, no-no.

Niamir A, Skidmore AK, Toxopeus AG, Muñoz AR, Real R (2011) Finessing atlas data for species distribution models. *Diversity and Distributions*, **17**, 1173-1185.

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

Imperfect detection

Braunisch V, Suchant R (- 2010) - Predicting species distributions based on incomplete survey data: the trade-off between precision and scale. *Ecography*, **33**, 826- 840.

Gu W, Swihart RK (2004) Absent or undetected? Effects of non-detection of species occurrence on wildlife–habitat models. *Biological Conservation*, **116**, 195-203.

Kéry M (2011) Towards the modelling of true species distributions. *Journal of Biogeography*, **38**, 617-618.

Le Lay G, Engler R, Franc E, Guisan A (2010) Prospective sampling based on model ensembles improves the detection of rare species. *Ecography*, **33**, 1015-1027.

Rota CT, Fletcher RJ, Evans JM, Hutto RL (2011) Does accounting for imperfect detection improve species distribution models? *Ecography*, **34**, 659-670.

Royle JA, Nichols JD, Kéry M (2005) Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, **110**, 353-359.

Selection of pseudo-absences and absences

Chefaoui RM, Lobo JM (2008) Assessing the effects of pseudo-absences on predictive distribution model performance. *Ecological Modelling*, **210**, 478-486.

Lobo JM, Jiménez-Valverde A, Hortal J (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103-114.

Václavík T, Meentemeyer RK (2009) Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions? *Ecological Modelling*, **220**, 3248-3258.

VanDerWal J, Shoo LP, Graham C, Williams SE (2009) Selecting pseudo-absence data for presence-only distribution modeling: How far should you stray from what you know? *Ecological Modelling*, **220**, 589-594.

Selection of geographical boundaries of study, background and distance constraints

Acevedo P, Jiménez-Valverde A, Lobo JM, Real R (2012) Delimiting the geographical background in species distribution modelling. *Journal of Biogeography*, in press.

Allouche O, Steinitz O, Rotem D, Rosenfeld A, Kadmon R (2008) Incorporating distance constraints into species distribution models. *Journal of Applied Ecology*, **45**, 599-609.

Anderson RP, 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.

Barbet-Massin M, Thuiller W, Jiguet F (- 2010) - How much do we overestimate future local extinction rates when restricting the range of occurrence data in climate suitability models? - *Ecography*, **33**, 878- 886.

Barve N, Barve V, Jiménez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J, Villalobos F (2011) The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, **11**, 1810-1819.

Sánchez-Fernández D, Lobo JM, Hernández-Manrique OL (2011) Species distribution models that do not incorporate global data misrepresent potential distributions: a case study using Iberian diving beetles. *Diversity and Distributions*, **17**, 163-171.

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

Addressing sample bias and sampling strategy

Phillips SJ, Dudik M, Elith J, Graham CH, Lehman A, Lathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181-197.

Albert CH, Yoccoz NG, Edwards TC, Graham CH, Zimmermann NE, Thuiller W (2010) Sampling in ecology and evolution ? bridging the gap between theory and practice. *Ecography*, **33**, 1028-1037.

Anderson RP, Gonzalez Jr. I (2011) Species-specific tuning increases robustness to sampling bias in models of species distributions: An implementation with Maxent. *Ecological Modelling*, **222**, 2796-2811.

De Cáceres M, Brotons L (2012) Calibration of hybrid species distribution models: the value of general-purpose vs. targeted monitoring data. *Diversity and Distributions*, no-no.

Edwards J, Thomas C., Cutler DR, Zimmermann NE, Geiser L, Moisen GG (2006) Effects of sample survey design on the accuracy of classification tree models in species distribution models. *Ecological Modelling*, **199**, 132-141.

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

McCarthy KP, Fletcher JR RJ, Rota CT, Hutto RL (2011) Predicting Species Distributions from Samples Collected along Roadsides; Predicción de la Distribución de Especies a Partir de Muestras Recolectadas a lo Largo de Carreteras. *Conservation Biology*, no-no.

Sheth SN, Lohmann LG, Distler T, Jiménez I (2011) Understanding bias in geographic range size estimates. *Global Ecology and Biogeography*, no-no.

Number of occurrences and prevalence in data

Feeley KJ, Silman MR (2011) Keep collecting: accurate species distribution modelling requires more collections than previously thought. *Diversity and Distributions*, no-no.

Hanberry BB, He HS, Dey DC (2012) Sample sizes and model comparison metrics for species distribution models. *Ecological Modelling*, **227**, 29-33.

Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773-785.

Mateo RG, Felicísimo ÁM, Muñoz J (2010) Effects of the number of presences on reliability and stability of MARS species distribution models: the importance of regional niche variation and ecological heterogeneity. *Journal of Vegetation Science*, no-no.

Pearson RG, Raxworthy CJ, Nakamura M, Townsend Peterson A (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102-117.

Wisz MS, Hijmans R, Li J, Peterson AT, Graham CH, Guisan A (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763-773.

Spatial autocorrelation in records

Cruse B, Liedloff AC, Wintle BA (2012) A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography*, no-no.

Ficetola GF, Manenti R, De Bernardi F, Padoa-Schioppa E (2012) Can patterns of spatial autocorrelation reveal population processes? An analysis with the fire salamander. *Ecography*, **35**, 693-703.

Marcer A, Pino J, Pons X, Brotons L (2012) Modelling invasive alien species distributions from digital biodiversity atlases. Model upscaling as a means of reconciling data at different scales. *Diversity and Distributions*, n/a-n/a.

Miller J, Franklin J (2002) Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. *Ecological Modelling*, **157**, 227-247

Václavík T, Kupfer JA, Meentemeyer RK (2011) Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography*, no-no.

Veloz SD (2009) Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. *Journal of Biogeography*, **36**, 2290-2299

ENVIRONMENTAL VARIABLES

Autocorrelation

Naimi B, Skidmore AK, Groen TA, Hamm NAS (2011) Spatial autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. *Journal of Biogeography*, no-no.

Climate versus/and other variables selection

Austin MP, Van Niel KP (2011) Impact of landscape predictors on climate change modelling of species distributions: a case study with *Eucalyptus fastigata* in southern New South Wales, Australia. *Journal of Biogeography*, **38**, 9-19.

Barbet-Massin M, Thuiller W, Jiguet F (2012) The fate of European breeding birds under climate, land-use and dispersal scenarios. *Global Change Biology*, **18**, 881-890.

Costa GC, Wolfe C, Shepard DB, Caldwell JP, Vitt LJ (2008) Detecting the influence of climatic variables on species distributions: a test using GIS niche-based models along a steep longitudinal environmental gradient. *Journal of Biogeography*, **35**, 637-646.

Lassueur T, Joost S, Randin CF (2006) Very high resolution digital elevation models: Do they improve models of plant species distribution? *Ecological Modelling*, **198**, 139-153.

Martínez B, Viejo RM, Carreño F, Aranda SC (2012) Habitat distribution models for intertidal seaweeds: responses to climatic and non-climatic drivers. *Journal of Biogeography*, no-no.

Morán-Ordóñez A, Suárez-Seoane S, Elith J, Calvo L, de Luis E (2012) Satellite surface reflectance improves habitat distribution mapping: a case study on heath and shrub formations in the Cantabrian Mountains (NW Spain). *Diversity and Distributions*, **18**, 588-602.

Schlaepfer DR, Lauenroth WK, Bradford JB (2012) Effects of ecohydrological variables on current and future ranges, local suitability patterns, and model accuracy in big sagebrush. *Ecography*, **35**, 374-384.

Thuiller W, Araújo MB, Lavorel S (2004) Do we need land-cover data to model species distributions in Europe? *Journal of Biogeography*, **31**, 353-361.

Triviño M, Thuiller W, Cabeza M, Hickler T, Araújo MB (2011) The Contribution of Vegetation and Landscape Configuration for Predicting Environmental Change Impacts on Iberian Birds. *PloS one*, **6**, e29373.

Zimmermann N, Edwards Jr T, Moisen G, Frescino T, Blackard J (2007) Remote sensing-based predictors improve distribution models of rare, early successional and broadleaf tree species in Utah. *Journal of Applied Ecology*, **44**, 1057-1067.

Temporal considerations

Bateman BL, VanDerWal J, Johnson CN (2011) Nice weather for bettings: using weather events, not climate means, in species distribution models. *Ecography*, no-no.

Blaum N, Schwager M, Wichmann MC, Rossmannith E (2011) Climate induced changes in matrix suitability explain gene flow in a fragmented landscape ? the effect of interannual rainfall variability. *Ecography*, no-no.

Early R, Sax DF (2011) Analysis of climate paths reveals potential limitations on species range shifts. *Ecology Letters*, **14**, 1125-1133.

Stanton JC, Pearson RG, Horning N, Ersts P, Akçakaya RH (2012) Combining static and dynamic variables in species distribution models under climate change. *Methods in Ecology and Evolution*, **3**, 349-357.

Zimmermann NE, Yoccoz NG, Edwards TC, Meier ES, Thuiller W, Guisan A, Schmatz DR, Pearman PB (2009) Climatic extremes improve predictions of spatial patterns of tree species. *Proceedings of the National Academy of Sciences*, **106**, 19723-19728.

Spatial resolution, number of predictors and scale

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

Randin CF, Engler R, Normand S et al. (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Global Change Biology*, **15**, 1557-1569.

Roura-Pascual N, Suarez AV, McNyset K et al. (2006) Niche differentiation and fine-scale projections for Argentine ants based on remotely sensed data. *Ecological Applications*, **16**, 1832-1841.

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

Synes NW, Osborne PE (2011) Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography*, **20**, 904-914.

TECHNIQUES

Comparisons

Bedia J, Busqué J, Gutiérrez JM (2011) Predicting plant species distribution across an alpine rangeland in northern Spain. A comparison of probabilistic methods. *Applied Vegetation Science*, no-no.

Brotos L, Thuiller W, Araújo MB, Hirzel AH (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, **27**, 437-448.

Elith* J, Graham* CH, Anderson RP et al. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129-151.

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

Graham CH, Hijmans RJ (2006) A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography*, **15**, 578-587.

Guisan A, Zimmermann NE, Elith J, Graham CH, Phillips S, Peterson AT (2007) What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs*, **77**, 615-630.

Leathwick JR, Elith J, Hastie T (2006) Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological Modelling*, **199**, 188-196.

Meynard CN, Quinn JF (2007) Predicting species distributions: a critical comparison of the most common statistical models using artificial species. *Journal of Biogeography*, **34**, 1455-1469.

Muñoz J, Felicísimo ÁM (2004) Comparison of statistical methods commonly used in predictive modelling. *Journal of Vegetation Science*, **15**, 285-292.

Pearson, R. G., Thuiller, W., Araújo, M. B., Martinez-Meyer, E., Brotos, L., McClean, C., Miles, L., Segurado, P., Dawson, T. P. and Lees, D. C. (2006), Model-based uncertainty in species range prediction. *Journal of Biogeography*, **33**, 1704–1711.

Thuiller W, Araújo MB, Lavorel S, Kenkel N (2003) Generalized models vs. classification tree analysis: Predicting spatial distributions of plant species at different scales. *Journal of Vegetation Science*, **14**, 669-680.

Model overparametrization / complexity

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

Ensemble modeling of CENM

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

Buisson L, Thuiller W, Casajus N, Lek S, Grenouillet G (2009) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, **16**, 1145-1157.

Grenouillet G, Buisson L, Casajus N, Lek S (2011) Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography*, **34**, 9-17.

Roura-Pascual N, Brotons L, Peterson AT, 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.

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

MODEL PERFORMANCE, EVALUATION, VALIDATION AND THRESHOLDS

Araújo MB, Pearson RG, Thuiller W, Erhard M (2005) Validation of species–climate impact models under climate change. *Global Change Biology*, **11**, 1504-1513

Austin MP, Belbin L, Meyers JA, Doherty MD, Luoto M (2006) Evaluation of statistical models used for predicting plant species distributions: Role of artificial data and theory. *Ecological Modelling*, **199**, 197-216.

Franklin J, Wejnert KE, Hathaway SA, Rochester CJ, Fisher RN (2009) Effect of species rarity on the accuracy of species distribution models for reptiles and amphibians in southern California. *Diversity and Distributions*, **15**, 167-177.

Hanberry BB, He HS, Dey DC (2012) Sample sizes and model comparison metrics for species distribution models. *Ecological Modelling*, **227**, 29-33.

Hanspach J, Kühn I, Schweiger O, Pompe S, Klotz S (2011) Geographical patterns in prediction errors of species distribution models. *Global Ecology and Biogeography*, no-no.

Liu C, White M, Newell G (2011) Measuring and comparing the accuracy of species distribution models with presence-absence data. *Ecography*, **34**, 232-243.

Marmion M, Luoto M, Heikkinen RK, Thuiller W (2009) The performance of state-of-the-art modelling techniques depends on geographical distribution of species. *Ecological Modelling*, **220**, 3512-3520.

McPherson JM, Jetz W, Rogers DJ (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 CN, Kaplan DM (2011) The effect of a gradual response to the environment on species distribution modeling performance. *Ecography*, no-no.

Smulders M, Nelson TA, Jelinski DE, Nielsen SE, Stenhouse GB (2010) A spatially explicit method for evaluating accuracy of species distribution models. *Diversity and Distributions*, no-no.

Syphard AD, Franklin J (2010) Species traits affect the performance of species distribution models for plants in southern California. *Journal of Vegetation Science*, **21**, 177-189.

Vaughan IP, Ormerod SJ (2005) The continuing challenges of testing species distribution models. *Journal of Applied Ecology*, **42**, 720-730.

Thresholds

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.

Fielding AH, Bell JF (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38.

Freeman EA, Moisen GG (2008) A comparison of the performance of threshold criteria for binary classification in terms of predicted prevalence and kappa. *Ecological Modelling*, **217**, 48-58.

Jiménez-Valverde A, Lobo JM (2007) Threshold criteria for conversion of probability of species presence to either–or presence–absence. *Acta Oecologica*, **31**, 361-369.

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385-393.

Moisen 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.

Nenzén H, Araújo M (2011) Choice of threshold alters projections of species range shifts under climate change. *Ecological Modelling*, **222**, 3346-3354.

AUC

Jiménez-Valverde A (2011) Insights into the area under the receiver operating characteristic curve (AUC) as a discrimination measure in species distribution modelling. *Global Ecology and Biogeography*, **21**, 498-507.

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

Peterson AT, Papeş M, Soberón J (2008) Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, **213**, 63-72.

PROJECTIONS IN SPACE AND TIME**Measuring extrapolation and novel situations**

Veloz SD, Williams JW, Blois JL, 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.

Webber BL, Yates CJ, Le Maitre DC et al. (2011) Modelling horses for novel climate courses: insights from projecting potential distributions of native and alien Australian acacias with correlative and mechanistic models. *Diversity and Distributions*, **17**, 978-1000.

Wenger SJ, Olden JD (2012) Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, **3**, 260-267.

Differences between climate change scenarios applied to CENM

Beaumont LJ, Hughes L, Pitman AJ (2008) Why is the choice of future climate scenarios for species distribution modelling important? *Ecology Letters*, **11**, 1135-1146.

Beaumont LJ, Pitman AJ, Poulsen M, Hughes L (2007) Where will species go? Incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, **13**, 1368-1385.

Garcia RA, Burgess ND, Cabeza M, Rahbek C, Araújo MB (2012) Exploring consensus in 21st century projections of climatically suitable areas for African vertebrates. *Global Change Biology*, **18**, 1253-1269.

Issues and strategies for projecting to past climate scenarios

Nogués-Bravo D (2009) Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, **18**, 521-531.

Pearman PB, Randin CF, Broennimann O et al. (2008) Prediction of plant species distributions across six millennia. *Ecology Letters*, **11**, 357-369.

Porch N (2010) Climate space, bioclimatic envelopes and coexistence methods for the reconstruction of past climates: a method using Australian beetles and significance for Quaternary reconstruction. *Quaternary Science Reviews*, **29**, 633-647.

Svenning J, Fløjgaard C, Marske KA, Nógues-Bravo D, Normand S (2011) Applications of species distribution modeling to paleobiology. *Quaternary Science Reviews*, **30**, 2930-2947.

Varela S, Lobo JM, Hortal J (2011) Using species distribution models in paleobiogeography: A matter of data, predictors and concepts. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **310**, 451-463

Transferability versus accuracy

Heikkinen RK, Marmion M, Luoto M (2011) Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography*, **35**, 276-288.

Syphard AD, Franklin J (2009) Differences in spatial predictions among species distribution modeling methods vary with species traits and environmental predictors. *Ecography*, **32**, 907-918.

Scaling

Borda-de-Água L, Borges PAV, Hubbell SP, Pereira HM (2011) Spatial scaling of species abundance distributions. *Ecography*, no-no.

ASSUMPTIONS AND BIOLOGICAL MEANING OF CENM**Dealing with local adaptations and sub-specific entities**

Benito Garzón M, Alía R, Robson TM, Zavala MA (2011) Intra-specific variability and plasticity influence potential tree species distributions under climate change. *Global Ecology and Biogeography*, **20**, 766-778.

Thompson GD, Robertson MP, Webber BL, Richardson DM, Le Roux JJ, Wilson JR (2011) Predicting the subspecific identity of invasive species using distribution models: *Acacia saligna* as an example. *Diversity and Distributions*, **17**, 1001-1014.

Equilibrium with climate

Araújo MB, Pearson RG (2005) Equilibrium of species? distributions with climate. *Ecography*, **28**, 693-695.

Leathwick JR (1998) Are New Zealand's *Nothofagus* species in equilibrium with their environment? *Journal of Vegetation Science*, **9**, 719-732.

Munguía M, Rahbek C, Diniz-Filho JAF, Araújo MB Equilibrium of Global Amphibian Species Distributions with Climate. *PLoS ONE*, e34420.

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

Species interactions and CENM

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

Banta JA, Ehrenreich IM, Gerard S, Chou L, Wilczek A, Schmitt J, Kover PX, Purugganan MD (2012) Climate envelope modelling reveals intraspecific relationships among flowering phenology, niche breadth and potential range size in *Arabidopsis thaliana*. *Ecology Letters*, n/a-n/a.

Heikkinen RK, Luoto M, Virkkala R, Pearson RG, Körber JH (2007) Biotic interactions improve prediction of boreal bird distributions at macroscales. *Global Ecology and Biogeography*, **16**, 754-763.

Bateman BL, VanDerWal J, Williams SE, Johnson CN (2012) Biotic interactions influence the projected distribution of a specialist mammal under climate change. *Diversity and Distributions*, n/a-n/a.

Heikkinen RK, Marmion M, Luoto M (2011) Does the interpolation accuracy of species distribution models come at the expense of transferability? *Ecography*, **35**, 276-288.

Hof AR, Jansson R, Nilsson C (2012) How biotic interactions may alter future predictions of species distributions: future threats to the persistence of the arctic fox in Fennoscandia. *Diversity and Distributions*, **18**, 554-562.

Meier ES, Kienast F, Pearman PB, Svenning J, Thuiller W, Araújo MB, Guisan A, Zimmermann NE (2010) Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, **33**, 1038-1048.

Meineri E, Skarpaas O, Vandvik V (2012) Modeling alpine plant distributions at the landscape scale: Do biotic interactions matter? *Ecological Modelling*, **231**, 1-10.

Pellissier L, Anne Bråthen K, Pottier J et al. (2010) Species distribution models reveal apparent competitive and facilitative effects of a dominant species on the distribution of tundra plants. *Ecography*, **33**, 1004-1014.

Wisz MS, Pottier J, Kissling WD et al. (2012) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological Reviews*, no-no.

Physiology and CENM

Austin MP, Smith TM, Niel KPV, Wellington AB (2009) Physiological responses and statistical models of the environmental niche: a comparative study of two co-occurring *Eucalyptus* species. *Journal of Ecology*, **97**, 496-507.

Buckley LB, Urban MC, Angilletta MJ, Crozier LG, Rissler LJ, Sears MW (2010) Can mechanism inform species? distribution models? *Ecology Letters*, **13**, 1041-1054.

Buckley LB, Waaser SA, MacLean HJ, Fox R (2011) Does including physiology improve species distribution model predictions of responses to recent climate change? *Ecology*, **92**, 2214-2221.

Helaouët P, Beaugrand G (2009) Physiology, Ecological Niches and Species Distribution. *Ecosystems*, **12**, 1235.

Keenan T, Serra J, Lloret F, Ninyerola M, Sabate S (2011) Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO₂ matters! *Global Change Biology*, **17**, 565-579.

Population dynamics and CENM

Brook BW, Akçakaya HR, Keith DA, Mace GM, Pearson RG, Araújo MB (2009) Integrating bioclimate with population models to improve forecasts of species extinctions under climate change. *Biology Letters*, **5**, 723-725.

Boulangéat I, Gravel D, Thuiller W (2012) Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. *Ecology Letters*, **15**, 584-593.

Buse J, Griebeler EM (2011) Incorporating classified dispersal assumptions in predictive distribution models - A case study with grasshoppers and bush-cricket. *Ecological Modelling*, **222**, 2130-2141.

Keith DA, Akçakaya HR, Thuiller W et al. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560-563.

Pagel J, Schurr FM (2011) Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography*, **21**, 293-304.

Regan HM, Syphard AD, Franklin J, Swab RM, Markovchick L, Flint AL, Flint LE, Zedler PH (2011) Evaluation of assisted colonization strategies under global change for a rare, fire-dependent plant. *Global Change Biology*, **18**, 936-947.

Schurr FM, Midgley GF, Rebelo AG, Reeves G, Poschlod P, Higgins SI (2007) Colonization and persistence ability explain the extent to which plant species fill their potential range. *Global Ecology and Biogeography*, **16**, 449-459.

Zhu K, Woodall CW, Clark JS (2011) Failure to migrate: lack of tree range expansion in response to climate change. *Global Change Biology*, **18**, 1042-1052.

Zurell D, Jeltsch F, Dormann CF, Schröder B (2009) Static species distribution models in dynamically changing systems: how good can predictions really be? *Ecography*, **32**, 733-744.

Modeling communities and or multispecies

Baselga A, Araújo MB (2009) Individualistic vs community modelling of species distributions under climate change. *Ecography*, **32**, 55-65.

Bittner T, Jaeschke A, Reineking B, Beierkuhnlein C (2011) Comparing modelling approaches at two levels of biological organisation ? Climate change impacts on selected Natura 2000 habitats. *Journal of Vegetation Science*, **22**, 699-710.

Ferrier S, Guisan A (2006) Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, **43**, 393-404.

Fitzpatrick MC, Sanders NJ, Ferrier S, Longino JT, Weiser MD, Dunn R (2011) Forecasting the future of biodiversity: a test of single- and multi-species models for ants in North America. *Ecography*, **34**, 836-847.

Godsoe W (2012) Are comparisons of species distribution models biased? Are they biologically meaningful? *Ecography*, **35**, 769-779. (Not directly related to multispecies distribution but to bias in direct comparisons among species distribution for multiple species).

Hallstan S, Johnson RK, Willén E, Grandin U (2012) Comparison of classification-then-modelling and species-by-species modelling for predicting lake phytoplankton assemblages. *Ecological Modelling*, **231**, 11-19.

Miller J, Franklin J (2002) Modeling the distribution of four vegetation alliances using generalized linear models and classification trees with spatial dependence. *Ecological Modelling*, **157**, 227-247.

Mokany K, Ferrier S (2011) Predicting impacts of climate change on biodiversity: a role for semi-mechanistic community-level modelling. *Diversity and Distributions*, **17**, 374-380.

Plant functional traits and species characteristics

Angert AL, Crozier LG, Rissler LJ, Gilman SE, Tewksbury JJ, Chunco AJ (2011) Do species? traits predict recent shifts at expanding range edges? *Ecology Letters*, no-no.

Douma JC, Witte JM, Aerts R, Bartholomeus RP, Ordoñez JC, Venterink HO, Wassen MJ, van Bodegom PM (2012) Towards a functional basis for predicting vegetation patterns; incorporating plant traits in habitat distribution models. *Ecography*, **35**, 294-305.

Hernandez PA, Graham CH, Master LL, Albert DL (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. *Ecography*, **29**, 773-785.

MAIN REVIEWS AND REFERENCE MATERIAL FOR CENM

Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, **33**, 1677-1688.

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

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

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

Elith J, Leathwick JR (2009) Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 677-697.

Franklin J (2010) *Mapping species distributions :spatial inference and prediction*. Cambridge University Press, Cambridge etc., 320 pp.

Franklin J (2010) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, **16**, 321-330.

Godsoe W (2010) I can't define the niche but I know it when I see it: a formal link between statistical theory and the ecological niche. *Oikos*, **119**, 53-60.

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

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.

Hampe A (2004) Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography*, **13**, 469-471. → Pearson RG, Dawson TP (2004) Bioclimate envelope models: what they detect and what they hide ? response to Hampe (2004). *Global Ecology and Biogeography*, **13**, 471-473.

Hirzel AH, Lay GL (2008) Habitat suitability modelling and niche theory. *Journal of Applied Ecology*, **45**, 1372-1381.

Hijmans RJ, Graham CH (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Global Change Biology*, **12**, 2272-2281.

Hirzel AH, 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.

Iverson LR, Prasad AM, Matthews SN, Peters MP (2011) Lessons learned while integrating habitat, dispersal, disturbance, and life-history traits into species habitat models under climate change. *Ecosystems*, **14**, 1005-1020.

Jiménez-Valverde A, Lobo JM, Hortal J (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, **14**, 885-890.

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

Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361-371.

Peterson AT, Soberon J, Pearson RG, Anderson RP, Martinez-Meyer E, Nakamura M, Araújo MB (2011) *Ecological Niches and geographic distributions*. 46, Princeton Univ Pr.

Sillero N (2011) What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling*, **222**, 1343-1346.

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

Thuiller W, Albert C, Araújo MB et al. (2008) Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology, Evolution and Systematics*, **9**, 137-152.

Zimmermann NE, Edwards TC, Graham CH, Pearman PB, Svenning J (2010) New trends in species distribution modelling. *Ecography*, **33**, 985-989.

Theoretical papers

Cassini MH (2011) Ecological principles of species distribution models: the habitat matching rule. *Journal of Biogeography*, no-no.

Chase JM, Leibold MA (2003) *Ecological niches :linking classical and contemporary approaches*. University of Chicago Press, Chicago etc., 212 pp.

Colwell RK, Rangel TF (2009) Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences*, **106**, 19651-19658.

Dawson TP, Jackson ST, House JI, Prentice IC, Mace GM (2011) Beyond Predictions: Biodiversity Conservation in a Changing Climate. *Science*, **332**, 53-58.

Hof C, Levinsky I, Araújo MB, Rahabek C (2011) Rethinking species' ability to cope with rapid climate change. *Global Change Biology*, **17**, 2987-2990.

Holt RD (2009) Bringing the Hutchinsonian niche into the 21st century: Ecological and evolutionary perspectives. *Proceedings of the National Academy of Sciences*, **106**, 19659-19665.

Jackson ST, Betancourt JL, Booth RK, Gray ST (2009) Ecology and the ratchet of events: Climate variability, niche dimensions, and species distributions. *Proceedings of the National Academy of Sciences*, **106**, 19685-19692.

Kylafis G, Loreau M (2011) Niche construction in the light of niche theory. *Ecology Letters*, **14**, 82-90.

Morin X, Lechowicz MJ (2008) Contemporary perspectives on the niche that can improve models of species range shifts under climate change. *Biology Letters*, **4**, 573-576.

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

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

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

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

Soberón JM (2010) Niche and area of distribution modeling: a population ecology perspective. *Ecography*, **33**, 159-167.

Soberón J, Nakamura M (2009) Niches and distributional areas: Concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences*, **106**, 19644-19650.

Wake DB, Hadly EA, Ackerly DD (2009) Biogeography, changing climates, and niche evolution. *Proceedings of the National Academy of Sciences*, **106**, 19631-19636.

Wiens JJ, Ackerly DD, Allen AP et al. (2010) Niche conservatism as an emerging principle in ecology and conservation biology. *Ecology Letters*, **13**, 1310-1324.

MAIN PATTERNS AND DESCRIPTIVE PAPERS

Baselga A, Lobo JM, Svenning J, Araújo MB (2012) Global patterns in the shape of species geographical ranges reveal range determinants. *Journal of Biogeography*, **39**, 760-771.

Benito Garzón M, Sánchez de Dios R, Sainz Ollero H (2008) Effects of climate change on the distribution of Iberian tree species. *Applied Vegetation Science*, **11**, 169-178.

Benito Garzón M, Sánchez de Dios R, Sainz Ollero H (2007) Predictive modelling of tree species distributions on the Iberian Peninsula during the Last Glacial Maximum and Mid-Holocene. *Ecography*, **30**, 120-134.

Berry PM, Dawson TP, Harrison PA, Pearson RG (2002) Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Global Ecology and Biogeography*, **11**, 453-462.

Garcia RA, Burgess ND, Cabeza M, Rahbek C, Araújo MB (2012) Exploring consensus in 21st century projections of climatically suitable areas for African vertebrates. *Global Change Biology*, **18**, 1253-1269.

Meier ES, Lischke H, Schmatz DR, Zimmermann NE (2011) Climate, competition and connectivity affect future migration and ranges of European trees. *Global Ecology and Biogeography*, no-no.

Nogués-Bravo D, Araújo MB, Errea MP, Martínez-Rica JP (2007) Exposure of global mountain systems to climate warming during the 21st Century. *Global Environmental Change*, **17**, 420-428.

Ruiz-Labourdette D, Nogués-Bravo D, Ollero HS, Schmitz MF, Pineda FD (2011) Forest composition in Mediterranean mountains is projected to shift along the entire elevational gradient under climate change. *Journal of Biogeography*, no-no.

Thomas CD, Cameron A, Green RE et al. (2004) Extinction risk from climate change. *Nature*, **427**, 145-148.

Thuiller W, Lavorel S, Araújo MB, Sykes MT, Prentice IC (2005) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 8245-8250.

APPENDIX 2

APPENDIX 2.1. Results from niche-based models. (NBM).

Table1.Variable Importance in NBM, by model and species studied. Variable importance is calculated as 1 minus the correlation between two models: the full model and one in which the target variable has been randomised. As the two models are highly correlated, resulting variable importance decreases.

SPECIES	MODEL	ACCUMULATED PRECIPITATION	MEAN ANNUAL TEMPERATURE	REAL SOLAR RADIATION
<i>Quercus ilex</i>	ANN	0.42	0.85	0.01
	CTA	0.38	0.88	0.06
	GAM	0.25	0.84	0.02
	GBM	0.32	0.87	0.01
	GLM	0.28	0.83	0.02
	MARS	0.39	0.79	0.04
	RF	0.60	0.83	0.26
<i>Pinus sylvestris</i>	ANN	0.82	0.30	0.04
	CTA	0.83	0.29	0.04
	GAM	0.77	0.180	0.03
	GBM	0.86	0.182	0.00
	GLM	0.78	0.24	0.03
	MARS	0.76	0.27	0.04
	RF	0.79	0.33	0.14

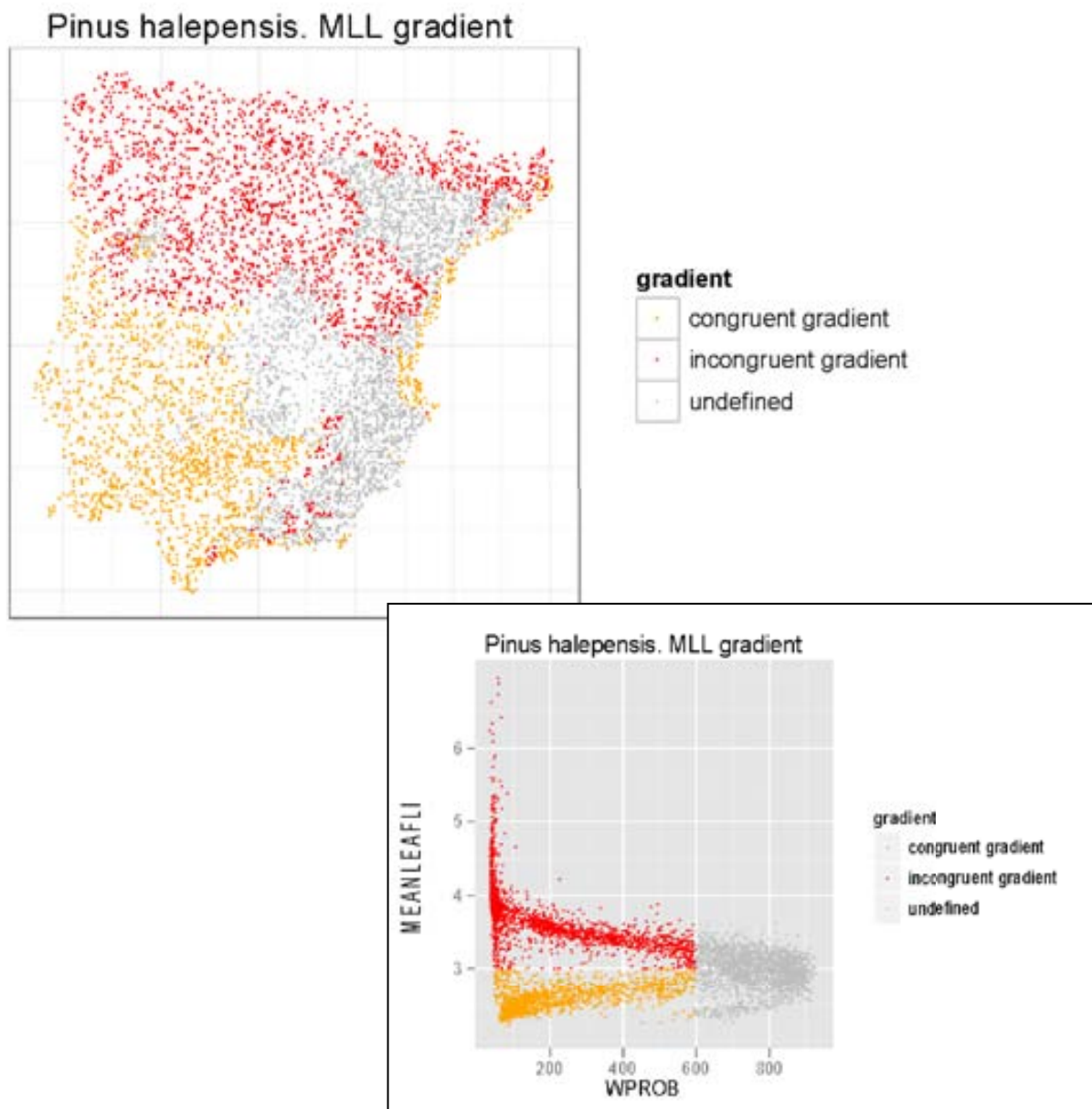
<i>Pinus halepensis</i>	ANN	0.67	0.51	0.04
	CTA	0.65	0.54	0.11
	GAM	0.50	0.52	0.04
	GBM	0.60	0.49	0.00
	GLM	0.58	0.49	0.06
	MARS	0.58	0.50	0.08
	RF	0.61	0.60	0.25

Table 2. Evaluation scores of Niche-based models by species using three different indexes: True Skill Statistic (TSS), Area under de curve (AUC) and Kappa statistic. Models: Artificial neural networks (ANN), Classification tree analysis (CTA), General additive models (GAM), Multivariate adaptive regression splines (MARS), Random forests (RF). See Table 1 in manuscript for further details.

SPECIES	MODEL	AUC	TSS	Kappa
<i>Quercus ilex</i>	ANN	0.769	0.428	0.391
	CTA	0.811	0.527	0.489
	GAM	0.824	0.524	0.498
	GBM	0.823	0.529	0.499
	GLM	0.819	0.514	0.488
	MARS	0.834	0.535	0.506
	RF	0.840	0.535	0.501
<i>Pinussylvestris</i>	ANN	0.863	0.629	0.599
	CTA	0.918	0.745	0.693
	GAM	0.929	0.740	0.695
	GBM	0.929	0.744	0.686
	GLM	0.925	0.726	0.690
	MARS	0.934	0.742	0.699
<i>Pinushalepensis</i>	RF	0.926	0.732	0.687
	ANN	0.806	0.484	0.443
	CTA	0.883	0.664	0.621
	GAM	0.901	0.669	0.605
	GBM	0.897	0.655	0.604
	GLM	0.898	0.660	0.593
	MARS	0.906	0.687	0.628
	RF	0.911	0.704	0.639

APPENDIX 2.2 Geography of mean leaf life (MLL) in simulations in *Pinus halepensis* .

Geography of the different physiological strategies for mean leaf life (MLL) in *Pinushalepensis*. Grey areas indicate the upper part of the suitability gradient where all simulated forests present similar values of net primary production (NPP). Orange areas and red areas indicate the a positive (congruent) or negative (incongruent) relationship between mean leaf life and weighted suitability



**APPENDIX 2.3. Analysis of variance results between
NPP and suitable/unsuitable virtual forests for
Quercus ilex and *Pinus sylvestris*.**

SPECIES		CONGRUENT RANGE		WHOLE RANGE	
		Kruskal-Wallis test	p	Kruskal-Wallis test	p
QUIL	ENS	602.77	4.18 E-133	113.794	1.45 E-26
	GLM	794.75	7.49 E-175	186.87	1.53 E-42
	RF	299.56	4.10 E-67	46.72	8.16 E-12
	ANN	1272.05	1.34 E-278	249.15	3.97 E-56
	CTA	485.36	1.46 E-107	51.56	6.96 E-13
	GAM	877.17	9.00 E-193	198.75	3.92 E-45
	MARS	747.49	1.41 E-164	132.38	1.23 E-30
PISY	ENS	790.19	7.33 E-174	790.19	7.33 E-174
	GLM	761.99	9.89 E-168	761.99	9.89 E-168
	RF	503.74	1.46 E-111	503.74	1.46 E-111
	ANN	489.95	1.46 E-108	489.95	1.46 E-108
	CTA	773.25	3.53 E-170	773.25	3.53 E-170
	GAM	759.59	3.29 E-167	759.59	3.29 E-167
	MARS	751.73	1.69 E-165	751.73	1.69 E-165

APPENDIX 3

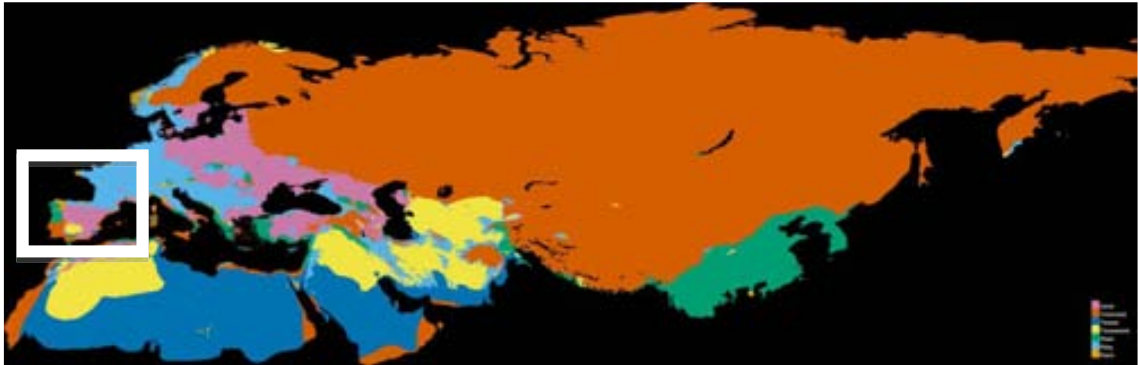
APPENDIX 3.1 Extrapolation in the environmental space.

Rhaponticum repens projections to the Palearctic realm. Different colors indicate novel climates where extrapolation occurs. Measures of extrapolation have been provided by Maxent extensions MESS and MoD (see Elith et al. 2010). Maps indicate the most dissimilar variable outside its training range. Most of projections to the Iberian Peninsula (square frame) show no extrapolation issues, except for some cases (European invasion). As shown, limiting dispersal increases extrapolation into projections.

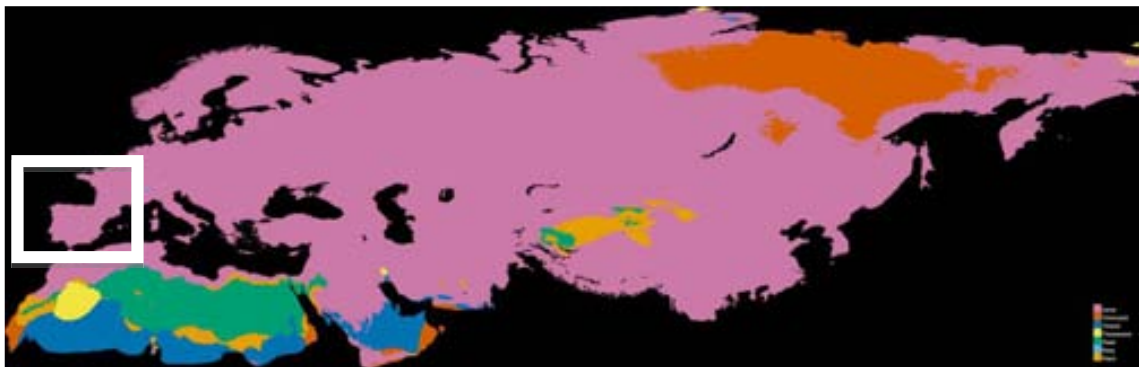
R.repens (Neotropic origin) - model maxent



R.repens (Neotropic origin) – model maxent dispersal constrained



R.repens (Nearctic origin) – model maxent



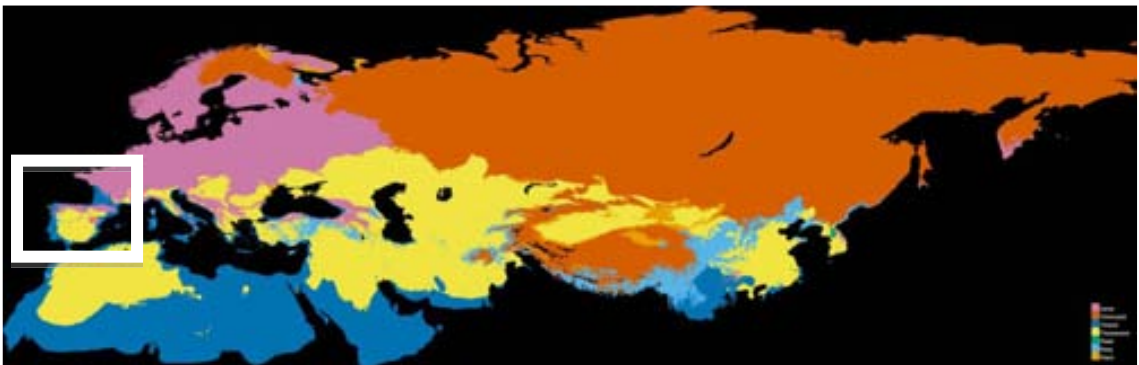
R.repens (Nearctic origin) – model maxent dispersal constrained



R.repens (Palearctic invasive origin) – model maxent



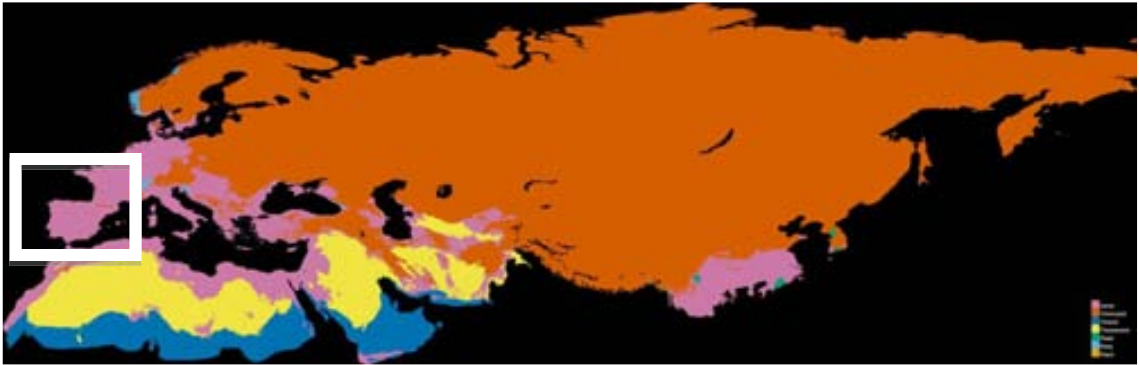
R.repens (Palearctic invasive origin) – model maxent dispersal constrained



R. repens (Australian origin) – model maxent



R. repens (Australian origin)– model maxent dispersal constrained



R. repens native range (Palearctic)– model maxent



R. repens native range (Palearctic)– model maxent dispersal constrained



APPENDIX 4

APPENDIX 4.1 Model descriptions

Species Distribution Models description (inspired by Thuiller et al. 2009 and Franklin 2010).

Random forests (RF): A machine-learning method – a combination of tree predictors in which each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. (Breiman, 2001)

Classification tree analysis (CTA): A classification method – a 50-fold cross-validation to select the best trade-off between the number of leaves of the tree and the explained deviance. (Breiman et al., 1984)

Multivariate adaptive regression splines (MARS): A non-parametric regression method, combining elements of CTA and GAM. (Friedman, 1991)

Generalized linear model (GLM): A regression method, with polynomial terms for which a stepwise procedure is used to select the most significant variables. (McCullagh & Nelder, 1989)

Generalized additive model (GAM): A regression method more flexible than GLM, we used a spline of 4 degrees of freedom and a stepwise procedure to select the most parsimonious model. (Hastie & Tibshirani, 1990)

Generalized Boosting Models (GBM): A method that fits a large tree of simple models, together aimed at giving a more robust estimate of the response. Based on Boosted Regression Tree algorithm. (Friedman, 2001)

Artificial neural networks (ANN): A machine-learning method, with the mean of three runs used to provide predictions and projections. (Ripley, 1996)

Flexible discriminant Analysis (FDA): A supervised classification method based on a mixture of normals obtain a density estimation of each class. (Hastie and Tibshirani, 1996)

REFERENCES

- Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). Classification and regression trees Chapman and Hall, New York.
- Breiman, L. (2001), Random Forests, Machine Learning 45(1), 5-32. Breiman, L (2002). Manual On Setting Up, Using, And Understanding Random Forests V3.1.
- Franklin, J. (2010). Mapping Species Distributions: Spatial Inference and Prediction. Cambridge Univ Pr, Cambridge, UK.
- Friedman, J.H. (1991). Multivariate Additive Regression Splines. Annals of Statistics.
- Friedman, J.H. (2001) Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29, 1189-1232
- Hastie, T.J. and Tibshirani, R. (1990). Generalized additive models Chapman and Hall, London.
- Hastie, T. and Tibshirani, R. (1996). Discriminant Analysis by Gaussian Mixtures. JRSSB.
- McCullagh, P. and Nelder, J.A. (1989). Generalized linear models. Chapman and Hall.
- Ripley, B.D. (1996) Pattern Recognition and Neural Networks. Cambridge.
- Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M. B. (2009). BIOMOD : A platform for ensemble forecasting of species distributions. Ecography, 32, 369-373.

APPENDIX 4.2 Model Accuracy

Models' accuracy measures by the True Skill Statistic (TSS, Allouche et al. 2006). Reported values indicated averaged TSS across model repetitions. Model abbreviation in Appendix 4.1

SPECIES	MODEL	TSS	TSS
		Cross validation	Total score
<i>Pinus balforuniana</i> (PIBA) N= 217	ANN	0.92	0.94
	CTA	0.90	0.93
	GAM	0.87	0.86
	GBM	0.94	0.95
	GLM	0.83	0.82
	MARS	0.89	0.89
	FDA	0.79	0.8
	RF	0.94	0.97
<i>Pinus coulterii</i> (PICO3) N=323	ANN	0.78	0.8
	CTA	0.76	0.82
	GAM	0.78	0.79
	GBM	0.82	0.85
	GLM	0.74	0.75
	MARS	0.75	0.76
	FDA	0.74	0.75
	RF	0.84	0.94
<i>Pinus muricata</i> (PIMU) N=65	ANN	0.68	0.77
	CTA	0.67	0.82
	GAM	0.66	0.67
	GBM	0.77	0.88
	GLM	0.63	0.64
	MARS	0.53	0.61
	FDA	0.52	0.59
	RF	0.74	0.91

APPENDICES

SPECIES	MODEL	TSS	TSS
		Cross validation	Total score
<i>Pinus sabiniana</i> (PISA2) N=2372	ANN	0.67	0.68
	CTA	0.7	0.75
	GAM	0.64	0.65
	GBM	0.67	0.69
	GLM	0.65	0.65
	MARS	0.65	0.65
	FDA	0.63	0.64
	RF	0.76	0.91
<i>Quercus douglasii</i> (QUDO) N= 2422	ANN	0.72	0.72
	CTA	0.77	0.8
	GAM	0.73	0.72
	GBM	0.75	0.76
	GLM	0.68	0.68
	MARS	0.73	0.73
	FDA	0.72	0.71
	RF	0.81	0.93
<i>Quercus dumosa</i> (QUDU) N= 83	ANN	0.68	0.75
	CTA	0.63	0.78
	GAM	0.65	0.65
	GBM	0.75	0.88
	GLM	0.67	0.67
	MARS	0.56	0.6
	FDA	0.39	0.43
	RF	0.7	0.89
<i>Quercus engelmannii</i> (QUEN) N = 36	ANN	0.77	0.84
	CTA	0.67	0.78
	GAM	0.67	0.69
	GBM	0.81	0.91
	GLM	0.75	0.72
	MARS	0.78	0.78
	FDA	0.66	0.67
	RF	0.71	0.89

SPECIES	MODEL	TSS	
		Cross validation	Total score
Quercus lobata (QULO) N = 699	ANN	0.65	0.66
	CTA	0.66	0.74
	GAM	0.65	0.64
	GBM	0.68	0.7
	GLM	0.62	0.6
	MARS	0.62	0.61
	FDA	0.60	0.6
	RF	0.72	0.9
Quercus sabiniana (QUSA2) N = 999	ANN	0.93	0.93
	CTA	0.92	0.93
	GAM	0.92	0.92
	GBM	0.93	0.94
	GLM	0.91	0.91
	MARS	0.90	0.89
	FDA	0.82	0.83
	RF	0.94	0.97
Quercus wislezenii (QUWI) N = 2763	ANN	0.54	0.53
	CTA	0.63	0.68
	GAM	0.57	0.57
	GBM	0.63	0.64
	GLM	0.57	0.57
	MARS	0.59	0.59
	FDA	0.57	0.57
	RF	0.68	0.89

APPENDIX 4.3 Metrics description

Broad species geographic attribute of conservation (range or current distribution). Each metric is provided with its formula and a description. In grey, relevant studies using the same or very similar metric/concept.

RANGE LEVEL	Conservation applied to potential suitable habitat whether currently occupied or not; range-wide ecosystem management perspective.
----------------	------------------------------------------------------------------------------------------------------------------------------------

Species range change (SRC)

$$SRC = A_{\text{Gained}} - A_{\text{Lost}} ;$$

where A= area

Differences in potential suitable area between two periods. Area Gained is the total area predicted suitable at time 2 but not time 1; Area Lost is predicted suitable at time 1 but not time 2.

Purpose: Indicates the degree of shrinkage or expansion of potential suitable habitat.

Range exposure to migration (REM)

Difference in habitat suitable area between full and null dispersal assumptions. Area of suitable habitat assuming full dispersal is the total area predicted suitable at time 2;

area assuming null (no) migration is the area of intersection of habitat predicted suitable at time 1 and time 2 (stable habitat)

$$REM = (A_{Full} - A_{Null}) / A_{Null};$$

where A= area

Purpose: Assess the degree of disparity between assumption of full migration and no migration. REM represents the potential role of migration on filling potential suitable habitat.

(Araujo & New, 2007 suggest bounding boxes between these assumptions; Svenning and Skov 2004 use current to potential distribution)

Range velocity (RV)

Differences the velocity of climatic exposure between leading edge (unsuitable cells in t_0 becoming suitable in t_1) and trailing edge (suitable cells in t_0 becoming unsuitable in t_1).

$$RV = v_{gained} - v_{lost} = \frac{\sum_{i=1, N} v_{i,j; \text{unsuitable} \rightarrow \text{suitable}}}{N} - \frac{\sum_{i=1, N} v_{i,j; \text{suitable} \rightarrow \text{unsuitable}}}{N}$$

where $v_{i,j}$ = bioclimatic velocity in cell $x_{i,j}$, and N is the number of cells. Velocity is calculated as follows:

$$V_{i,j} = \frac{T_{i,j}}{S_{i,j}} ;$$

Where $T_{i,j}$ is the temporal gradient of probabilities and $S_{i,j}$ is the spatial gradient of probabilities in cell $x_{i,j}$.

Gradients are defined as:

$$T_{i,j} = \int_{t_0}^{t_1} \frac{dP_{i,j}}{dt} = \frac{P_{i,j;t_1} - P_{i,j;t_0}}{t_1 - t_0} ;$$

$$S_{i,j} = \sqrt{(S_i^2 + S_j^2)} ;$$

where S_i is spatial gradient in direction i and S_j is spatial gradient in direction j (see figure)

Cell position scheme (focal cell in bold):

$X_{i-1,j+1}$ (a)	$X_{i,j+1}$ (b)	$X_{i+1,j+1}$ (c)
$X_{i-1,j}$ (d)	$X_{i,j}$ (e)	$X_{i+1,j}$ (f)
$X_{i-1,j-1}$ (g)	$X_{i,j-1}$ (h)	$X_{i+1,j-1}$ (i)

$$S_i = \frac{dPe}{dx} = \frac{(Pc+2Pf+Pi) - (Pa + 2Pd + Pg)}{8 \times cellsize}$$

$$S_j = \frac{dPe}{dx} = \frac{(Pc+2Pf+Pi) - (Pa + 2Pd + Pg)}{8 \times cellsize}$$

Purpose: To assess the balance between the velocity between trailing edge and leading edge in order to determine which one is faster in acquiring or losing suitable conditions.

Range spatial fragmentation (RSF)

Number of discrete habitat patches at a given time period.

Purpose: To assess spatial fragmentation of potential suitable habitats.

(Opdam&Wascher, 2004)

Range Spatial Aggregation (RSA)

Percent of total suitable habitat occupied by the largest (suitable) patch at a given time period.

$$LPI = (A_{\text{LargestPatch}} / A_{\text{suitable}}) * 100$$

A= area

Purpose: To assess spatial aggregation of potential suitable habitats.
(Opdam&Wascher, 2004)

PLOT LEVEL	Conservation applied to current known distribution (forest plots); local management strategies and <i>in situ</i> conservation.
---------------	------------------------------------------------------------------------------------------------------------------------------------

Forest migration effort (FME)

Average distance of the most climatically suitable route for current forest locations (plot) to reach a suitable patch. Cost resistance surface is determined by probabilities of presence.

$$FME = \frac{\sum_{i=0,N} \int_{t_0}^{t_1} dx_{i,j(P)} (C)}{N}$$

where i is forest plot, j is nearest suitable patch determined by the probability function (P) and dx is the distance between the two points as a function of the cost resistance surface C (product of linear distance to suitable patches and inverse of suitability for each cell of the grid).

Purpose: to incorporate a surrogate metric of meta-population persistence and potential connectivity processes.

(Skov and Svenning 2004, use tree cover for resistance , Wang et al. 2008 found significant relationship between habitat suitability resistance and gene-flow)

Forest climate-site exposure (FCE)

Percentage of current forest locations (plots) changing from suitable in t_0 to unsuitable conditions in t_1 .

$$FCE = (N_{t_1_suitable \rightarrow t_2_unsuitable} / N) \times 100;$$

where N is the number of plots.

Purpose: to assess the magnitude of current forests exposed to new environmental conditions. (Thomas et al. 2004; Thuiller et al. 2005)

Forest climatic velocity (FCV)

Mean bioclimatic velocity of forests plots in decline.

$$FCV = \frac{\sum_{i=1, N} v_i}{N}$$

Where velocity (v) in plot (i) is assigned as the velocity in matching cell $x_{i,j}$ as in computed as in the range velocity metric (see above), and N indicating plots of decreasing suitability in time.

Purpose: to assess the speed of change in exposure of current forests.

REFERENCES

- Araújo, M. B. & New, M. (2007). Ensemble forecasting of species distributions, 22, 42-47.
- Opdam, P. & Wascher, D. (2004). Climate change meets habitat fragmentation: Linking landscape and biogeographical scale levels in research and conservation. *Biol.Conserv.*, 117, 285-297.
- Skov, F. & Svenning, J. (2004). Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography*, 27, 366-380.
- Svenning, J. & Skov, F. (2004). Limited filling of the potential range in European tree species. *Ecol.Lett.*, 7, 565-573.
- Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F. N., de Siqueira, M. F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A. S., Midgley, G. F., Miles, L., Ortega-Huerta, M., Townsend Peterson, A., Phillips, O. L. & Williams, S. E. (2004). Extinction risk from climate change. *Nature*, 427, 145-148.
- Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T. & Prentice, I. C. (2005). Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 8245-8250.
- Wang, Y. H., Yang, K. C., Bridgman, C. L. & Lin, L. K. (2008). Habitat suitability modelling to correlate gene flow with landscape connectivity. *Landscape Ecol.*, 23, 989-1000.

APPLYING CORRELATIVE ECOLOGICAL NICHE MODELS TO GLOBAL CHANGE STUDIES

The usefulness of climate-species distribution approaches

Doctorat en Ciències Ambientals

Tesi doctoral realitzada per

J. M. Serra Díaz

Tesi doctoral dirigida per

Miquel Ninyerola Casals



Universitat Autònoma de Barcelona

Departament de Biologia Animal, de Biologia Vegetal i d'Ecologia



UAB

Universitat Autònoma de Barcelona

2012