bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Local patch occupancy dynamics match the distribution of European trees at continental scale

Arnaud Guyennon[1], Björn Reineking[1], Jonas Dahlgren[2], Gerald Kändler[3], Aleksi Lehtonen[4], Sophia Ratcliffe[5], Paloma Ruiz-Benito[6,7], Miguel A. Zavala[7], and Georges Kunstler[1]

[1] *Univ. Grenoble Alpes, Irstea, LESSEM, 38000 Grenoble, France*
[2] *Swedish University of Agricultural Sciences, , Umeå, 90183 Sweden*
[3] *Forest Research Institute Baden-Wurttemberg, , Freiburg, 79100 Germany*
[4] *Natural Resources Institute Finland (Luke), , Latokartanonkaari 9 FI-00790 Helsinki Finland*
[5] *Department of Systematic Botany and Functional Biodiversity, University of Leipzig, Johannisallee 21-23, 04103 Leipzig, Germany*
[6] *Departamento de Biología y Geología, Física y Química Inorgánica, Escuela Superior de Ciencias Experimentales y Tecnología, Universidad Rey Juan Carlos, C/ Tulipán s/n, 28933, Móstoles (Madrid), Spain*
[8] *Grupo de Ecologıa y Restauracion Forestal, Departamento de Ciencias de la Vida, Universidad de Alcala, Edificio de Ciencias, Campus Universitario, 28805 Alcala de Henares, Madrid, Spain*

Abstract

Processes driving current tree species distribution are still largely debated, and large uncertainties about processes and scales at play remain. Attempts to relate environmental factors and population related metrics have shown mixed results, and the link between tree species range and their population dynamics are unclear. In this context, we would like to test the hypotheses that metapopulation processes drive range limits. We defined for each of the 17 tree species analyzed a species distribution model (SDM) relating environmental variables to their current distribution. We then test whether colonization and extinction probabilities are related to probability of presence estimated by the SDMs. Harmonized national forest inventories across Europe, that includes several censuses, were used to calibrate a dynamic patch occupancy model based on the probability of presence derived from SDMs. Results show for most species that colonization probabilities increased with increasing probability of presence, and conversely extinction probabilities decreased. Using these results, we evaluated whether the equilibrium probability of presence derived form the patch occupancy model match the observed species distribution.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1 Introduction

The vast majority of species have restricted geographical range distributions (Holt & Keitt, 2000). Understanding the factors determining these ranges is fundamental to have insights on future species redistribution in the face of climate change. Species distribution is thought to be tightly connected to its ability to cope with local abiotic conditions and thus species' niche (Pulliam, 2000; Thuiller, Lavorel, & Araújo, 2005; Soberón, 2007). This view underpins most statistical species distribution models (SDM) that relate species presence with local environmental conditions. These models have been extensively used in recent years and provided very detailed descriptions of species environmental requirements based on occurrence data. They provide, however, very little indication on how species distribution arise from population dynamics. This is surprising because we start to have a rich theoretical understanding of how population dynamics controls species distribution and species range (Holt, Keitt, Lewis, Maurer, & Taper, 2005). The first and most classical view is that species ranges match the environmental conditions where birth rates exceed mortality rates (i.e. where the rate of population growth is above 1). We have very few field tests of this assumption, and available results are contradictory. For instance, Thuiller et al. (2014) demonstrated that major demographic parameters of European tree species were not strongly correlated to the probability of occurrence derived from SDMs. In contrast, a limited number of transplant experiments beyond species range have shown that population growth rate or some demographic rates tend to decline beyond the distributional edge but these experiments are generally limited to few species and cover a very restricted sample of the species distribution (Hargreaves, Samis, & Eckert, 2013; Lee-Yaw et al., 2016). In a complementary way, several studies did not find strong support for the abundance center hypothesis (Brown, 1984) which propose that abundance should be higher in the center of the distribution (Murphy, VanDerWal, & Lovett-Doust, 2006; Sagarin, Gaines, & Gaylord, 2006). A second class of theoretical explanations rely on the idea that stochastic events increase local extinction risks outside of the species ranges and limits it presence (Holt et al., 2005). Variability in demographic rates due to environmental variability, as well as stochastic demographic events, may thus have a considerable impact on population persistence (Boyce, Haridas, Lee, & Group, 2006; Ovaskainen & Meerson, 2010). Field tests of these mechanisms are extremely rare. Csergő et al. (2017) used detailed demographic data for plant species (including trees) to test relationships between climate suitability, derived from a SDM, and several detailed population metrics including time to quasi-extinction, stochastic population growth rate or transient population dynamics. Results did not strongly support the idea that climate suitability is linked to demographic stochasticity.

The last class of mechanisms underlying species ranges is a regional equilibrium between colonization/extinction and dispersal (Holt et al., 2005; Holt & Keitt, 2000). This last category relates to the meta-population paradigm and proposes that species ranges arise from the gradient of three variables: the extinction rate, the colonization rate, and the habitat structure. An important simplification is that the dynamic of the local population is ignored. This is equivalent to assuming that local population processes are much faster than regional dynamics (extinction/colonization) and we can separate them (Drechsler & Wissel, 19997). Such models ignore the details of the population dynamics but rather focus on patch occupancy dynamics (extinction and colonization events). Few studies have focused on these processes for tree species (see Purves, Zavala, Ogle, Prieto, & Be-

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

nayas, 2007; García-Valdes, Zavala, Araújo, & Purves, 2013; García-Valdés, Gotelli, Zavala, Purves, & Araújo, 2015; Talluto, Boulangeat, Vissault, Thuiller, & Gravel, 2017) and we still lack studies that explore this mechanism at the European scale.

Here we propose to analyze species extinction and colonization events at the local scale for the main tree species in Europe across their entire continental distribution using national forest inventories data for more than 80 000 plots. To describe environmental species distribution we use species probability of presence estimated with ensemble SDM to the EU-Forest data set (Mauri, Strona, & San-Miguel-Ayanz, 2017), a large scale presence absence dataset over continental Europe. We then analyze the relationship between the probability of occurrence derived from SDMs and the extinction/colonization events to test the following hypothesis:

- Extinction rates increase when the SDM probability of presence decrease.
- Colonization rates decrease when the SDM probability of presence decrease.
- The equilibrium probability of occurrence predicted by simple patch occupancy models based on the estimated extinction/colonization rates matched the SDM estimated probability of presence.

This allows us to evaluate the relative importance of extinction and colonization in driving the range for different species. Finally, we discuss the implication of our results for the degree of equilibrium with current climate of these species.

2 Materials and Methods

For our analysis we first gathered data on tree population local extinction and colonization events from national forest inventories plots. We extracted information on local relative probability of presence of each species based on maps at a 1 km scale accounting for forest cover at a finer scale (JRC maps, see below). This provides a surrogate of the seed source for the colonization. Next, to describe in a simple way the position of each plot inside the species range, we computed the probability of presence with SDM models based on the EU-Forest data. Subsequently, we estimated the extinction and colonization rates from two observations of occupancy data via a spatially inhomogeneous Markov chain. Finally, we derived the probability of presence at equilibrium based on the estimated extinction and colonization rates.

2.1 NFI datasets

To calibrate our model, we required information on the presence/absence of each species at two different dates, over a large geographical area to cover, which, as far as possible, encompasses the entire species distribution. We used a database of tree data from the National Forest Inventories (NFI) of Finland, France, Germany, Spain, Sweden, compiled as part of the FunDivEurope project (http://www.fundiveurope.eu, Baeten et al. (2013)).

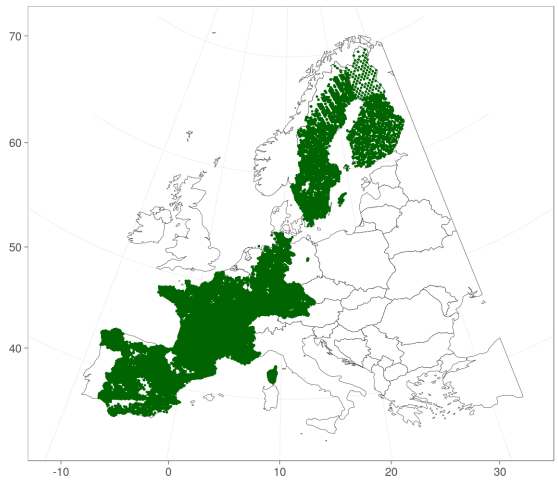
bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Inventory protocols differed between NFIs (see Appendix 1 for a detailed description of each survey protocol). Surveys were conducted within a circular plot with a fixed radius or in concentric subplots with different radius and minimum diameter at breast height (DBH) for all NFIs except Germany, where an angle-count method (basal area factor of 4 m^2ha^{-1}) was used. Because the DBH thresholds for trees to be recorded varied between the inventories, we only included trees with a DBH of 10 cm or greater. For each NFI, except France, two inventory surveys were conducted with a variable time interval (from 4 to 16 years, see figure 1b). The French inventory is based on single surveys where the growth of alive trees (based on a short core) and approximate date of death of dead trees are estimated and can be used to reconstruct the stand structure five years before the census, making it comparable with permanent plot data.

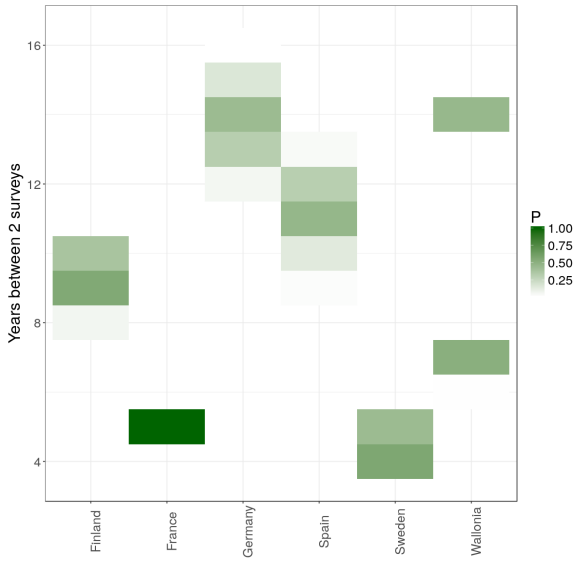
To avoid influences of management on the extinction and colonization processes, we discarded plots where a management effect was reported between surveys. This led to a selection of 80 157 plots with 173 species. Within these species we selected the most abundant species (the cumulative basal area of species retained represented more than 95 % of the total basal area) and excluded exotic species and species for which maps of local presence were not available. This resulted in a selection of 20 species.

For each plot, a species was considered as present when at least one tree was observed. The succession of two surveys allowed then to deduce state transitions ($0 \rightarrow 1$ for local colonization, $1 \rightarrow 0$ for local extinction). Since several protocols are based on concentric circular plots with varying DBH thresholds, a newly observed tree might not be a recruited tree: its DBH during the first census was above 10 cm, but it was not recorded due to the higher DBH threshold for the subplot. We used a species specific growth model to estimate the probability that a new tree (present only in the second census) had a former DBH below 10 cm. The growth model was built as a generalized linear model using an aridity index, the sum growing degree days and tree DBH as explanatory variables. We thus considered a plot as colonized if the probability that the largest newly observed tree had a former DBH below 10 cm greater or equal to 0.5, otherwise the species was considered as present at both censuses ($1 \rightarrow 1$). This correction had a strong impact on the Spanish and German inventories, significantly reducing the number of colonization events. We decided to exclude from further analysis species with less than 10 events for extinction or colonization (*i.e.* *Quercus suber*, *Pinus pinea* and *Acer pseudoplatanus*), resulting in a final selection of 17 species.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



(a) Locations (fuzzied at 500 m) of the plots used.



(b) Distribution of number of years between subsequent surveys by country.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2.2 Joint research center relative probability of presence

The density of NFI plots is too low to accurately describe the local abundance of trees that can disperse seeds into the plot. Distances between NFI plots are about 1 km or above whereas most dispersal events occur in less than 100 m from the seed source (see for example Nathan, Safriel, and Imanuel (2001), Bullock et al. (2017)). To represent seed input into a plot, we thus used the species' relative probability of presence (hereafter P_{JRC}) in the corresponding 1 km cell produced by the Joint Research Center (JRC), see San-Miguel-Ayanz, de Rigo, Caudullo, Houston Durrant, and Mauri (2016). Each map estimates the relative frequency of the species based on datasets of field observations as represented in the Forest Information System for Europe (FISE), which integrates National Forest inventories, BioSoil and Forest Focus. The presence/absence data are assimilated at a spatial resolution of 1 km based on multiple smoothing kernels of varying dimension. Independent estimations of forest cover extracted from the Pan-European Forest Type Map 2006 (FTYP2006, http://forest.jrc.ec.europa.eu/forest-mapping/forest-type-map) are then used to rescale the species frequency by the cover of broadleaved forest, coniferous forest or other non-forest categories based on 25 m x 25 m pixels.

2.3 SDM

We estimated species probability of presence (P_{occ}) on each NFI plot with ensemble species distribution models fitted the EU-Forest data set (Mauri et al., 2017) that provides species presence/absence on a 1 km grid. For each grid point we extracted mean annual temperature, precipitation of wettest quarter, temperature and precipitation seasonality from CHELSA climatologies (Karger et al., 2017), pH from SoilGrid (Hengl et al., 2017), and aridity index and actual evapo-transpiration from CGIAR-CSI (Trabucco & Zomer, 2010). Then we fitted ensemble SDM models with BIOMOD2 (Thuiller, Lafourcade, Engler, & Araújo, 2009) using four different models (GAM, GLM, GBM, and Random Forest). Based on this ensemble model we estimated species probability of presence on each NFI plot for all analyzed species. Details on the evaluation on the predictive power of the SDM are provided in the Appendix 2 (see Figure 1 with performance scores of SDM for each species based on True Skill Statistic, TSS and Area Under the Curve of the Receiver Operating Characteristic, AUC).

2.4 Patch occupancy transition model

The patch occupancy model is a spatially inhomogeneous Markov chain, the state vector being the patch occupancy of the N plots $X(t)$ at time t and the probability of transition between the two time successive patch occupancy patterns is:

$$P[\mathbf{X}(t+1)|\mathbf{X}(t)] = \prod_{i=1}^{N} \left[\begin{array}{llll} 1-C_i(t) & if\ X_i(t)=0 & and & X_i(t+1)=0 \\ C_i(t) & if\ X_i(t)=0 & and & X_i(t+1)=1 \\ E_i(t) & if\ X_i(t)=1 & and & X_i(t+1)=0 \\ 1-E_i(t) & if\ X_i(t)=1 & and & X_i(t+1)=1 \end{array} \right. \quad (1)$$

with N being the total number of plots observed, E_i the extinction probability in plot i, C_i the colonization probability in plot i.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The extinction probability (E) of a species in a plot only depends on the local grid conditions, i.e. the probability of presence derived from the SDM (P_{occ}). The colonization probability (C) is divided into two contributions: recruitment probability (R) which depends on the SDM probability of presence (P_{occ}), and seed availability (S). The recruitment probability R is the probability for at least one tree to reach 10 cm between two protocols, therefore integrates both growth and survival of the early stages of trees.

Colonization probability is simply expressed as the product of R and S, and each patch with a non-null seed availability can experience a colonization event. An explicit representation of seed availability via dispersal mechanisms was beyond the scope of this work, with seed dispersion being largely limited to short distances compared to the 1 km grid resolution of the field data (Nathan et al., 2001). From a formal perspective, seed availability from surroundings sources is equivalent to a kernel density estimate to estimate local prevalence, here taken from JRC relative probability of presence (P_{JRC}). This probability is is solely based on distance kernels, and the proportion of tree cover within the 1 km cell around the plot. We have chosen this variable because it summarizes a very large amount of data on a European scale and can be considered as an strong indicator of the proportion of adjacent plots in which the species is present within a 1 km patch.

Recruitment (R) and extinction (E) probabilities were related to the SDM probability of presence with the following equations:

$$\begin{aligned} &logit(E_i) = \alpha + \beta * P_{occi} \\ &logit(R_i) = \gamma + \delta * P_{occi} \\ &C_i = R_i * S_i, with \;\; S_i = P_{JRCi} \end{aligned} \tag{2}$$

To take into account the difference in protocol between countries, intercept parameters (α and γ) are country specific.

Because the time interval between two censuses may vary across plots (between 4 and 15 years), we standardized the parameters to a 5 years sampling interval, the probabilitiy of an recruitment/extinction was computed as:

$$P(Event) = 1 - (1 - P(Event_{5years}))^{\frac{n}{5}} \tag{3}$$

with n being the number of years between the two censuses.

2.5 Calibration of the model

For each species, extinction and colonization parameters were estimated separately using a Metropolis Hastings Monte Carlo sampling algorithm, with priors following a Cauchy distribution (Gelman, Jakulin, Pittau, & Su, 2008; Ghosh, Li, & Mitra, 2018) using JAGS (Plummer, 2003). Convergence was checked by evaluating whether the Gelman-Rubin convergence statistic was below 1.1, as recommended by Brooks and Gelman (1998), using 4 chains.

2.6 Probability of presence at equilibrium

Given estimates of recruitment and extinction probabilities, one can derive a probability of presence at equilibrium (hereafter P_{eq}). This probability of presence would be equivalent to the proportion

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of patches occupied in a 1 km JRC cell.

The equilibrium can be defined in two ways: (1) For an open system with a fixed seed source S, where the probability of presence in the grid cell is a function of extinction, recruitment and the value of seed source. In this case there is no feedback of the colonization and extinction on the seed source. (2) For a closed system, we can compute the probability of presence as the proportion of suitable patches occupied within each grid cell, with a feedback of the colonization and extinction processes on the seed source. In this case, an extinction probability exceeding the colonization probability would lead to a species absence.

Both types of equilibrium can be related to the same type of model as both formulations rely on the same equation:

$$\frac{dp}{dt} = C * (1 - p) - E * p = 0 \tag{4}$$

with p the proportion of suitable patches occupied.

The difference between the two types of equilibrium corresponds to different formulations of C. In the first formulation, S is constant over time and $C = R \times S$, while in the second formulation, S varies with the local SDM probability of presence and $C = p * R$. These two alternative formulations lead to the following equilibria:

- (1) when we consider a fixed seed source, and compute the equilibrium state for each plot: $P_{eq} = \frac{R \times S}{R \times S + E}$,
- (2) when we consider that the seed source is linked to the proportion of occupied patches within each 1 km grid cell, then the proportion of suitable occupied patches is $P_{eq} = (1 - \frac{E}{R})$.

For both formulations, we studied the relative impact of extinction and colonization on the equilibrium state by fixing one of the probabilities to its mean and letting the other vary based on our estimated slope of response to the SDM probability of occurrence. We can also compute the probability at equilibrium, letting both extinction and colonization vary with the SDM probability of occurrence. In the first model we can also set the fixed seed source to one (no dispersal limitation) or let the fixed seed source vary with the SDM probability of presence based on their correlation.

3 Results

3.1 Colonization/Extinction dependencies

For most species, recruitment probability increased with the SDM probability of presence, and extinction probability decreased with the SDM probability of presence, see figure 2.

When considering recruitment probability, slopes were significantly positive (confidence interval not intercepting zero) except for *Quercus petraea*, *Abies alba* and *Picea abies* (see table 1). For these species, the model fit was poor as indicated by their low value of ΔDIC (difference of deviance information criterion – DIC – between the model and a null model). No species showed a significant negative slope, i.e. a significant increase of recruitment probability when reaching the edge of the climatic niche.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

For extinction, slopes were significantly negative (confidence interval not including zero), for all species except *Pinus nigra*, *Quercus pubescens*, *Pinus pinaster* and *Pinus halepensis* (see table 2). These four species had the lowest value of ΔDIC. Only *Pinus pinaster* showed a posterior distribution of the slope that can be considered strictly positive.

Some species, such as *Populus tremula*, *Acer campestre* and *Alnus glutinosa*, exhibited a very broad posterior for the slope parameters. This is probably related to the small range of probabilities of occurrence for these species.

Model performances according to the True Skill Statistics (TSS, see (Allouche, Tsoar, & Kadmon, 2006)) varied from good (> 0.5)) to poor (< 0.3, (Araújo, Alagador, Cabeza, Nogués-Bravo, & Thuiller, 2011)) depending on the species and process (tables 1 and 2). Recruitment models showed good performance for all species but *Quercus pubescens*. Extinction models showed poor performance for *Pinus halepensis, Pinus pinaster, Quercus ilex, Populus tremula* and *Quercus pubescens* and good scores for *Quercus petraea* and *Quercus pyrenaica*. Model scores were not related to the number of observations.

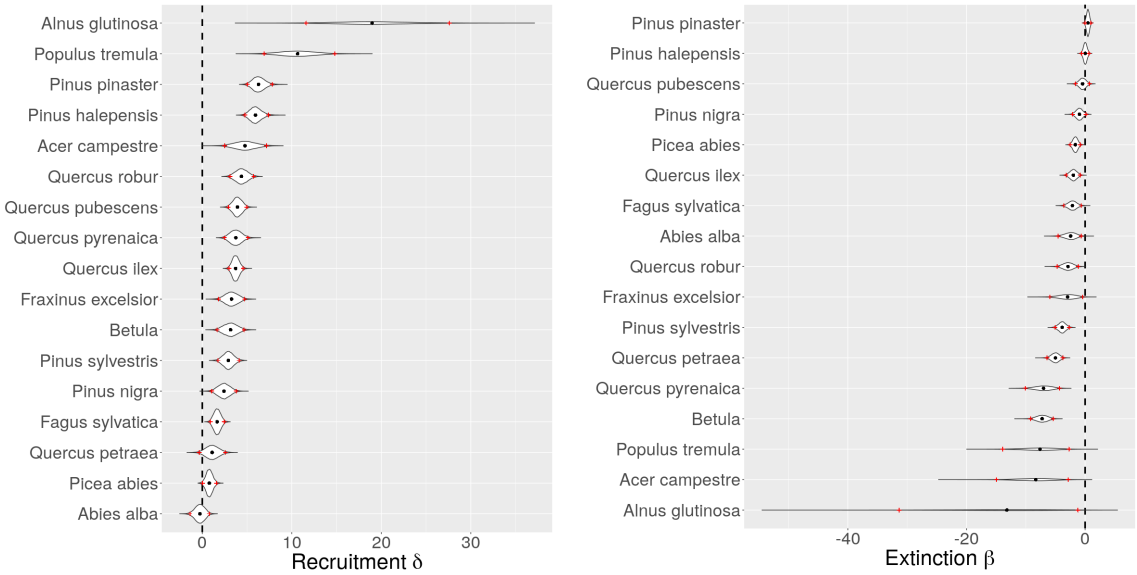


Figure 2: Posterior distribution of linear parameters of recruitment (left, δ) and extinction (right, β). Black points are posterior medians, red crosses indicate the 95th percentile intervals.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Species	Nb of events	Median β	95% Interval β	TSS	Δ DIC
Pinus sylvestris	149	2.9	1.8/4.1	0.60	-25.5
Picea abies	231	0.8	0.0/1.6	0.73	-2.7
Fagus sylvatica	160	1.7	0.9/2.4	0.62	-17.0
Quercus robur	103	4.4	3.1/5.7	0.65	-46.6
Quercus petraea	48	1.1	-0.3/2.6	0.66	1.3
Pinus pinaster	66	6.3	5.0/7.8	0.82	-118.9
Quercus ilex	351	3.7	3.0/4.6	0.64	-122.7
Pinus nigra	48	2.4	1.1/3.8	0.74	-11.4
Abies alba	53	-0.3	-1.4/0.8	0.63	1.3
Pinus halepensis	50	5.9	4.8/7.4	0.90	-136.2
Quercus pubescens	87	3.9	2.9/5.0	0.76	-61.4
Betula	265	3.2	1.7/4.7	0.58	-20.5
Fraxinus excelsior	121	3.3	1.8/4.7	0.53	-17.4
Quercus pyrenaica	85	3.7	2.5/5.1	0.85	-40.0
Alnus glutinosa	47	19.0	11.6/27.6	0.64	-25.8
Populus tremula	73	10.7	6.9/14.8	0.61	-32.7
Acer campestre	97	4.8	2.5/7.2	0.69	-15.7

Table 1: Estimated recruitment parameter β of the slope of response to P_{occ} per species and its 95% confidence interval (see Materials and Methods for details on the model). ΔDIC is the difference of deviance information criterion – DIC – between the model and a null model. TSS is the True Skill Statistics. Nb of events is the number of colonization events.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Species	Nb of events	Median α	95 % Interval α	TSS	Δ DIC
Pinus sylvestris	151	-3.9	-5.1/-2.7	0.27	-43.6
Picea abies	98	-1.7	-2.5/-0.8	0.31	-13.2
Fagus sylvatica	44	-2.1	-3.6/-0.8	0.31	-7.5
Quercus robur	70	-2.9	-4.5/-1.3	0.38	-11.4
Quercus petraea	101	-5.0	-6.4/-3.7	0.44	-67.8
Pinus pinaster	227	0.5	0/1.0	0.18	-1.7
Quercus ilex	50	-2.0	-3.2/-0.8	0.24	-9.7
Pinus nigra	67	-1.0	-2.1/0.1	0.15	-1.0
Abies alba	18	-2.5	-4.6/-0.6	0.35	-6.4
Pinus halepensis	117	0.0	-0.6/0.7	0.10	2.0
Quercus pubescens	40	-0.4	-1.6/0.8	0.19	1.1
Betula	188	-7.3	-9.3/-5.4	0.25	-66.0
Fraxinus excelsior	40	-2.9	-5.8/-0.4	0.30	-4.5
Quercus pyrenaica	36	-7.0	-10.2/-4.4	0.43	-36.6
Alnus glutinosa	21	-13.0	-32.1/-1.1	0.49	-6.8
Populus tremula	75	-7.7	-13.7/-2.6	0.28	-12.0
Acer campestre	30	-8.4	-15.3/-3.0	0.33	-12.5

Table 2: Estimated extinction parameter α of the slope of response to P_{occ} per species and its 95% confidence interval (see Materials and Methods for details on the model). ΔDIC is the difference of deviance information criterion – DIC – between the model and a null model. TSS is the True Skill Statistics. Nb of events is the number of extinction events.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We are interested in the relative effect of SDM probability of presence (P_{occ}) on extinction/recruitment. Since the ranges of P_{occ} differ across species, and the link function is not linear, the slope is not sufficient to evaluate the magnitude of the effect of the SDM probability of presence on recruitment and extinction. We thus also computed the relative contribution of the probability of occurrence to extinction and recruitment (figure 3) as the difference between the probability of extinction (colonization) at the low *vs.* high end of its respective range (respectively 5 and 95 % percentiles of P_{occ}).

For most species, the relative contribution of the probability of occurrence was higher for recruitment than for extinction, (most species are above the diagonal in the figure 3, particularly for *Quercus ilex*, *Quercus pubescens* and *Pinus halepebsis*). There is however important variation between species.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

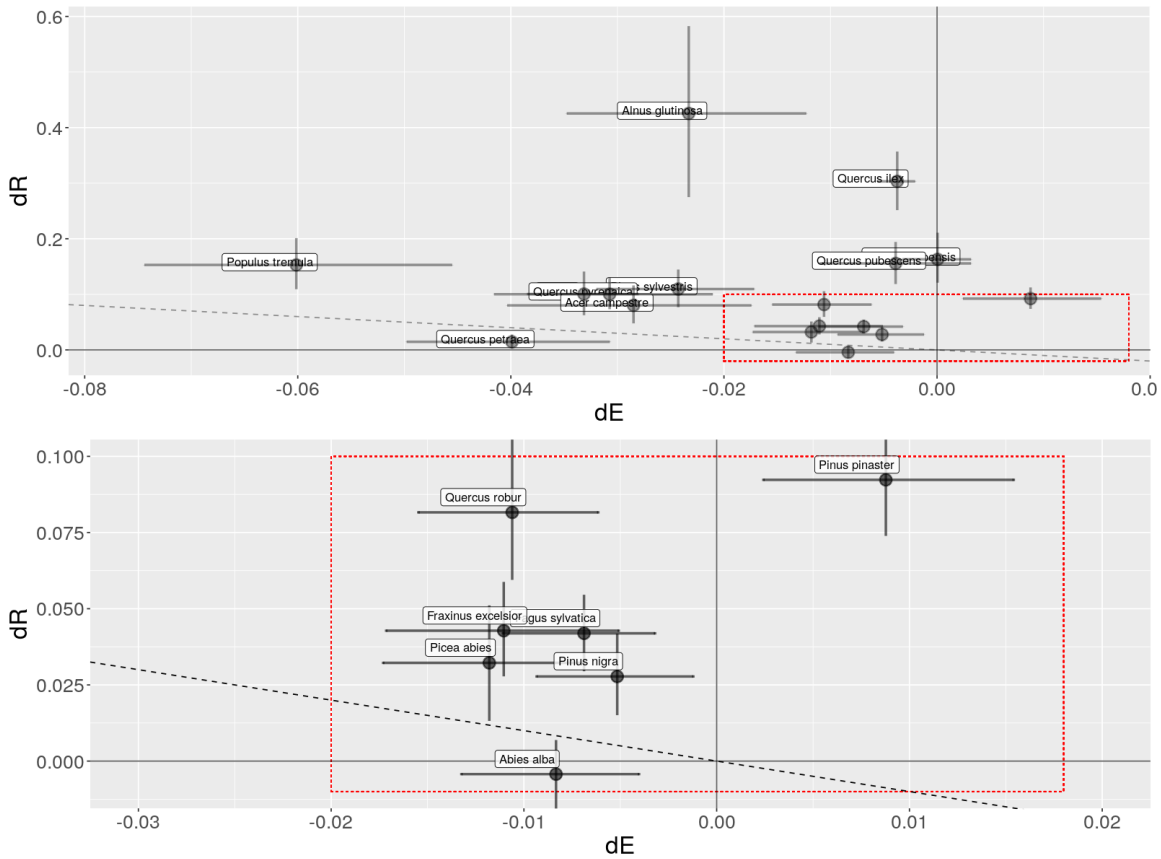


Figure 3: Relative contribution of the probability of occurrence to recruitment/extinction probabilities (dR and dE, respectively). For each species, dE and dR are calculated as the differences at high P_{occ} (95th centile) and low P_{occ} (5th centile). Bottom figure is a zoom of the top figure, illustrated with red lines. Negative dE means a higher extinction rate at low probability of presence; positive dR means a lower colonization at low probability of presence. On both plots dashed line is the bisecting line.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.2 Equilibrium

We observed a positive relationship between P_{eq} and P_{occ} for the four different methods of computing the probability of equilibrium with a fixed seed source (see figure 4). The relationship was always positive when we accounted for the variation of both recruitment and extinction probability, but there is no clear evidence that either recruitment or extinction probability had a stronger contribution to P_{eq}. When the seed source was fixed to one, P_{eq} showed less variation and overall overestimated P_{occ} (see figure 4). P_{eq} is above 0.5 when colonization probability exceeds extinction probability, which was almost always the case when dispersal limitation was not included (S set to one). Notable exceptions were *Pinus sylvestris* and *Pinus pinaster* for which extinction at low P_{occ} was substantially higher than the recruitment probability.

The impact of seed variations was variable between species. Including variations in S had a strong impact on P_{eq} variations only for three species (*Quercus ilex, Fagus sylvatica*, and *Picea abies*). When the seed source (S) was set to the observed JRC value in each cell, P_{eq} was very close to P_{occ} but generally still slightly overestimated P_{occ}. For four species the mean of P_{eq} was lower than P_{occ} for at least some value of P_{occ} (*Pinus nigra, Pinus pinaster, Pinus sylvestris*, and *Quercus robur*).

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

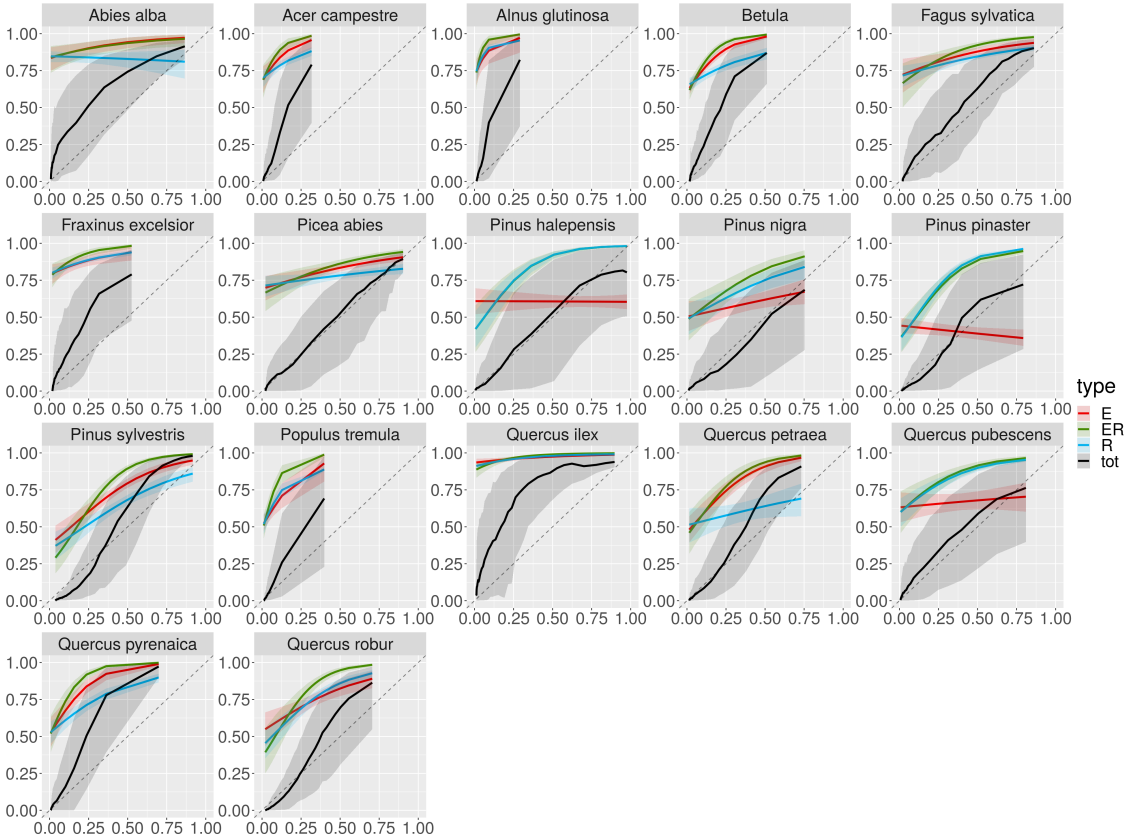


Figure 4: Equilibrium probability of occupancy against SDM probability of presence, calculated with seed source set to one and varying extinction (E in red), varying recruitment (R in blue), both (ER in green). S was computed with the seed source based on its correlation with P_{occ}.Black line (tot) is the equilibrium when considering all variations (E, R and S).

The second formulation of the equilibrium for a closed system leads also to a positive relationship between P_{eq} and P_{occ} for the three different ways of computing the equilibrium (E, R, and ER) (see figure 5). In contrast with the fixed source method, we here observed that at low values of P_{occ}, P_{eq} was below P_{occ}. In addition, P_{eq} could reach zero, meaning an absence of the species and quicly rise to higher values with a steeper slope than in the fixed seed source method.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

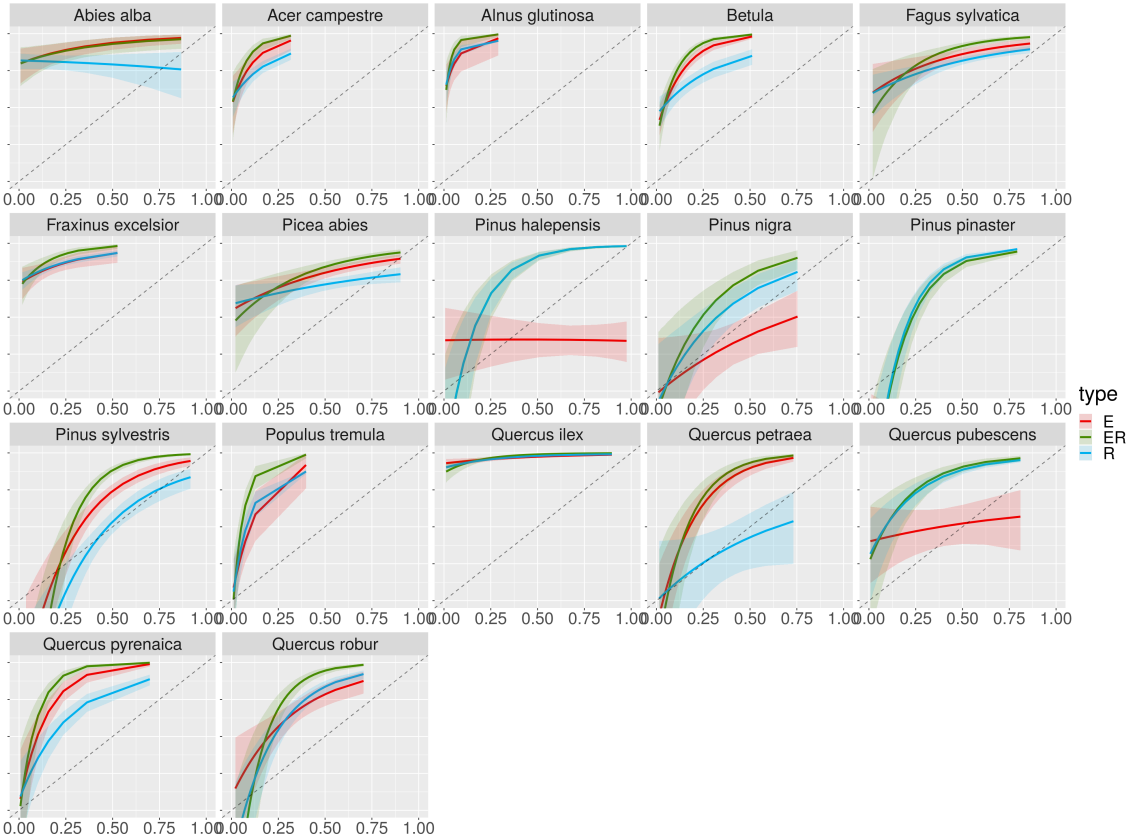


Figure 5: Equilibrium probability of occupancy against SDM probability of presence, calculated with closed formulation and varying extinction (E), varying recruitment (R), both (ER).

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4 Discussion

Overall, our results showed that on the edge of species climatic niche associated with a low probability of presence, the recruitment probability decreases and the extinction probability increases. We also found that the equilibrium occupancy probabilities, derived from recruitment and colonization probabilities are positively correlated with the observed probability of presence.

4.1 Variation of extinction and recruitment probability within species range

Our results reveal that species environmental distribution as defined by the SDM is positively related with colonization events and negatively related with extinction events. Only few species showed non significant links between the SDM probability of presence and extinction or recruitment (*Quercus pubescens,Pinus pinaster, Abies alba* and *Pinus halepensis*). Only *Pinus pinaster* showed an opposite response for extinction to the SDM probability of presence. The lack of relationships or opposite relationships for species belonging to the genus *Pinus* might be related to their intensive management and recent introduction in several regions (particularly in the case of *Pinus pinaster*). Overall, recruitment variability is higher than extinction variability, but, results did not show clear pattern with species ecological strategies, such as shade tolerance as previously reported (Thuiller et al., 2014). Indeed, assuming shade tolerant species are better competitors, we could have expected them to show a closer relationship between SDMs and the recruitment probability as this link would be less blurred by competition.

For the other species our results are compatible with Holt et al. (2005), and the environmental gradient as defined by the SDM probability of presence appears to be a valuable descriptor of the patch occupancy dynamics. Other studies on forest ecosystems have already shown that tree population extinction-colonization probabilities respond to the climate niche of species. This was generally done with models of patch occupancy (Talluto et al., 2017; Purves et al., 2007; García-Valdes et al., 2013) using polynomial functions of climatic variables (such as temperature and aridity). A key difference between our model and previous patch occupancy models in forests is that we did not use directly climatic variables but instead used the SDM probability of presence as a descriptor of species environmental niches. Given the low number of colonization or extinction events that we have despite using a very large number of plots using an SDM to summarize the species climatic niche might be more powerful than fitting complex multivariate responses to climatic variables. A similar approach has also been developed for birds in Britain by Araújo, Williams, and Fuller (2002).

The links of the SDM probability of presence to extinction/recruitment rates seems stronger than links to demographic rates reported in previous studies (Thuiller et al., 2014; Csergő et al., 2017). This might indicate that links are more easy to capture with integrative meta-population metrics than with detailed population level metrics.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4.2 Equilibrium *vs.* non-equilibrium of species distribution

If the distribution of a species is currently in equilibrium, we would expect a close relationship between the SDM probability of presence and the probability of presence computed at equilibrium. Observing a positive correlation, however, does not rule out that there may be a lag between the equilibrium and the current distribution due to climate change. Simulations (not presented here) show that as soon as the lag is not too large there should still be a positive correlation between the two probabilities. Results from our model tend to show that current species distribution is not far from a potential equilibrium. However the match between the extinction-colonization derived probability of presence and the SDM probability of presence is not exact. When we accounted for variation of both recruitment and extinction probability we found a positive relationship between the predicted probability of presence and the SDM probability of presence for both the fixed seed source and the closed models. We can see a good overall agreement between P_{eq} and P_{occ}. For most species, we observe a higher P_{eq} than P_{occ}, which would result in a higher prevalence in the area where the species is already present. We observed underestimation of P_{occ} only with the closed models.

The idea that each species is in current equilibrium with the environment has been criticized by Svenning and Skov (2004), based on the idea that most European tree species do not fully fill their potential ranges. This situation could be the result of a post-glacial migration lag as illustrated in Svenning, Normand, and Kageyama (2008). The lag would strongly affect *Abies alba*, the *Pinus* genus and the *Quercus* genus. This argument has however been partly contradicted by previous SDM results (Araújo & Pearson, 2005) and large dispersal rates found based on pollen records (Giesecke, Brewer, Finsinger, Leydet, & Bradshaw, 2017). Interestingly, we found a weak response of colonization and extinction to SDM probability of presence for *Abies alba*, a species with a recorded slow expansion rate (Giesecke et al., 2017).

4.2.1 Future redistribution

The match between P_{eq} and P_{occ}, for low values of P_{occ}, is important to understand future changes in species distribution as this corresponds to the edge of the niche. For instance, we could expect a reduction of the species range when P_{eq} is notably lower than P_{occ}, as observed for *Pinus nigra, Pinus pinaster, Pinus sylvestris*, and *Quercus petraea*. These results for *Pinus sylvestris* are coherent with simulations by Cheaib et al. (2012) and the sensitivity of the species to warming (Reich & Oleksyn, 2008) which may reduce its southern range.

Given the model formulation, we cannot conclude on the direction of species range expansion since the model does not differentiate the hot and cold edge, however it is likely that a higher P_{eq} than P_{occ} would result in range increases, as for example predicted for *Alnus glutinosa*, *Fraxinus excelsior* or *Quercus ilex* which is in agreement with previous results by Ruiz-Benito et al. (2017) in which climate change was expected to favour angiosperm traits/strategies.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

4.3 Different model types to compute the equilibrium, and dispersal

The two models used to compute equilibrium are two extreme cases of dispersal ability. In the open system formulation, dispersal is fixed arbitrarily to the current neighborhood conditions and does not change with time, while in the closed source model, dispersal is limited to the 1 km pixel. Reality probably lies in between these two cases, and dispersal kernel formulation as well as seed production are keys to go beyond these two simplified models (Nathan & Muller-Landau, 2000). The first formulation, based on current seed availability estimation, is more optimistic, as illustrated by the tendency to overestimate P_{occ} by P_{eq}. In the second formulation, equilibrium depend only on current environmental conditions (as summarized by the SDM), and the absence of dispersal between plots leads to strong variability. As illustrated by theoretical work, dispersal distance, through kernel formulation and its parameters have direct impact on potential population size (Law, Murrell, & Dieckmann, 2003) and its environmental responses (Snyder & Chesson, 2004). In addition, species range filling is closely linked to dispersal ability (Schurr et al., 2007).

4.4 Limitations of patch occupancy models

Several factors might have contributed to limit our ability to estimate the links between SDM probability of presence and extinction/recruitment probability. First, the NFI data do not provide a perfect knowledge of the absence/presence at the plot scale. With protocols that are based on concentric circular plots, we might miss the presence of tree larger than 10 cm DBH. We partially corrected this issue by accounting for the probability that the tree was below 10 cm at the first census with a growth model, but this approach is not perfect and the dataset probably might still contain colonization events that are not true colonizations but observation errors. Conversely, we might have wrongfully excluded some colonization events for trees with extreme growth. Secondly, even though we excluded plots where any management intervention (independantly of the forest management objectives) was recorded between the two censuses, management may still have had an impact. For example, the atypical responses of *Pinus pinaster* and *Pinus nigra* can be linked to the fact that these species are highly managed, and have been introduced in several regions. More broadly, the characterization of species climatice niche by an SDM suffer from being parametrized on a current distribution, and therefore account for a realized niche rather than a fundamental niche (Pulliam, 2000, e.g.). The SDMs may then fail to capture some climatic dependencies.

Different studies on patch occupancy models calibrated with NFI data have tried to formally include dispersal in the model (Purves et al., 2007; García-Valdés et al., 2015). García-Valdes et al. (2013) tried to infer the parameters of the dispersal kernel based on the Spanish forest inventory data. We considered that actual knowledge on the potential seed source surrounding a plot is insufficient to draw mechanistic conclusions on seed dispersal. Field studies show that mean distance of seed dispersal are short for most tree species (Nathan et al., 2001; Bullock et al., 2017; Cain, Milligan, & Strand, 2000), therefore direct dispersal between plots should be restricted to extremely rare events (distance > 1 km) and the tail of the kernel distribution. It is thus very unlikely that these models were really estimating a dispersal kernel (as indicated by the very large mean dispersal distance inferred) but rather captured a degree of spatial auto-correlation in the species distribution and

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the recruitment process. Here, we use a relative probability of presence which takes into account observed presence/absence and smoothing kernels as well as fine scale forest cover maps (building on the approach of Talluto et al. (2017)). We believe that if we want to include dispersal kernels in the model it is better to use external information on the shape and parameters of the dispersal kernel (see Schurr et al. (2007) or Schurr et al. (2012) for example).

Finally, our approach does not include biotic interactions and disturbances that might influence population extinction and recruitment probabilities (Case, Holt, McPeek, & Keitt, 2005; Svenning et al., 2014; Liang, Duveneck, Gustafson, Serra-Diaz, & Thompson, 2018). Given the small number of colonization and extinction events, a reliable estimate of these processes with our data seems unrealistic.

5 Conclusions

We tested how extinction/colonization rates were related to the variation of the probability of presence derived from SDM. For most tree species, we found an increase of recruitment probability and a decrease of extinction probability with the SDM probability of presence. Both relationships are in agreement with the hypothesis of range limits linked to extinction/colonization rate variations along an environmental gradient (Holt & Keitt, 2000). We did not see evidence for differences in the relative contributions to equilibrium probability of occupancy of extinction and recruitment. Models of species distribution build on the meta-population framework offer the promise to provide additional insights into the processes controlling species distribution and future changes in response to climate change.

6 Acknowledgments

This work was funded under EU FP7 ERA-NET Sumforest 2016 through the call "Sustainable forests for the society of the future" (project REFORCE), with the ANR as national funding agency (grant ANR-16-SUMF-0002). The authors are grateful to Christian Wirth for access to the harmonised inventory data. MAZ and PRB were supported by project FUNDIVER (Spanish Ministry of Economy and Competitiveness, MINECO, Spain, CGL2015-69186-C2-2-R) The NFI data synthesis was conducted within the FunDivEUROPE project funded by the European Union's Seventh Programme (FP7/2007–2013) under grant agreement No. 265171. We thank the MAGRAMA, the Johann Heinrich von Thunen-Institut, the Natural Resources Institute Finland (LUKE), the Swedish University of Agricultural Sciences, and the French Forest Inventory (IGN) for making NFI data available.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (tss). *Journal of Applied Ecology*, *43*(6), 1223–1232.

Araújo, M. B., Alagador, D., Cabeza, M., Nogués-Bravo, D., & Thuiller, W. (2011). Climate change threatens european conservation areas. *Ecology letters*, *14*(5), 484–492.

Araújo, M. B. & Pearson, R. G. (2005). Equilibrium of species' distributions with climate. *Ecography*, *28*(5), 693–695.

Araújo, M. B., Williams, P. H., & Fuller, R. J. (2002). Dynamics of extinction and the selection of nature reserves. *Proceedings of the Royal Society of London B: Biological Sciences*, *269*(1504), 1971–1980.

Baeten, L., Verheyen, K., Wirth, C., Bruelheide, H., Bussotti, F., Finér, L., ... Allan, E., et al. (2013). A novel comparative research platform designed to determine the functional significance of tree species diversity in european forests. *Perspectives in Plant Ecology, Evolution and Systematics*, *15*(5), 281–291.

Boyce, M. S., Haridas, C. V., Lee, C. T., & Group, N. S. D. W. (2006). Demography in an increasingly variable world. *Trends in Ecology & Evolution*, *21*(3), 141–148.

Brooks, S. P. & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*(4), 434–455.

Brown, J. H. (1984). On the relationship between abundance and distribution of species. *The American Naturalist*, *124*(2), 255–279.

Bullock, J. M., Mallada González, L., Tamme, R., Götzenberger, L., White, S. M., Pärtel, M., & Hooftman, D. A. (2017). A synthesis of empirical plant dispersal kernels. *Journal of Ecology*, *105*(1), 6–19.

Cain, M. L., Milligan, B. G., & Strand, A. E. (2000). Long-distance seed dispersal in plant populations. *American Journal of Botany*, *87*(9), 1217–1227.

Case, T. J., Holt, R. D., McPeek, M. A., & Keitt, T. H. (2005). The community context of species' borders: Ecological and evolutionary perspectives. *Oikos*, *108*(1), 28–46.

Cheaib, A., Badeau, V., Boe, J., Chuine, I., Delire, C., Dufrêne, E., ... Leadley, P. (2012). Climate change impacts on tree ranges: Model intercomparison facilitates understanding and quantification of uncertainty. *Ecology Letters*, *15*(6), 533–544.

Csergő, A. M., Salguero-Gómez, R., Broennimann, O., Coutts, S. R., Guisan, A., Angert, A. L., ... Buckley, Y. M. (2017). Less favourable climates constrain demographic strategies in plants. *Ecology letters*, *20*(8), 969–980.

Drechsler, M. & Wissel, C. (19997). Separability of local and regional dynamics in metapopulations. *Theoretical Population Biology*, *51*, 9–21.

García-Valdés, R., Gotelli, N. J., Zavala, M. A., Purves, D. W., & Araújo, M. B. (2015). Effects of climate, species interactions, and dispersal on decadal colonization and extinction rates of Iberian tree species. *Ecological Modelling*, *309-310*, 118–127.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

García-Valdes, R., Zavala, M. A., Araújo, M. B., & Purves, D. W. (2013). Chasing a moving target: Projecting climate change-induced shifts in non-equilibrial tree species distributions. *Journal of Ecology*, *101*, 441–453.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, *2*(4), 1360–1383.

Ghosh, J., Li, Y., & Mitra, R. (2018). On the use of cauchy prior distributions for bayesian logistic regression. *Bayesian Analysis*, *13*(2), 359–383.

Giesecke, T., Brewer, S., Finsinger, W., Leydet, M., & Bradshaw, R. H. (2017). Patterns and dynamics of european vegetation change over the last 15,000 years. *Journal of Biogeography*, *44*(7), 1441–1456.

Hargreaves, A. L., Samis, K. E., & Eckert, C. G. (2013). Are species' range limits simply niche limits writ large? a review of transplant experiments beyond the range. *The American Naturalist*, *183*(2), 157–173.

Hengl, T., de Jesus, J. M., Heuvelink, G. B., Gonzalez, M. R., Kilibarda, M., Blagotić, A., ... Mantel, S. (2017). Soilgrids250m: Global gridded soil information based on machine learning. *PLoS one*, *12*(2).

Holt, R. D. & Keitt, T. H. (2000). Alternative causes for range limits: A metapopulation perspective. *Ecology Letters*, *3*, 41–47.

Holt, R. D., Keitt, T. H., Lewis, M. A., Maurer, B. A., & Taper, M. L. (2005). Theoretical models of species' borders: Single species approaches. *Oikos*, *108*(1), 18–27.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific data*, *4*, 170122.

Law, R., Murrell, D. J., & Dieckmann, U. (2003). Population growth in space and time: Spatial logistic equations. *Ecology*, *84*(1), 252–262.

Lee-Yaw, J. A., Kharouba, H. M., Bontrager, M., Mahony, C., Csergő, A. M., Noreen, A. M., ... Angert, A. L. (2016). A synthesis of transplant experiments and ecological niche models suggests that range limits are often niche limits. *Ecology letters*, *19*(6), 710–722.

Liang, Y., Duveneck, M. J., Gustafson, E. J., Serra-Diaz, J. M., & Thompson, J. R. (2018). How disturbance, competition, and dispersal interact to prevent tree range boundaries from keeping pace with climate change. *Global Change Biology*, *24*(1), e335–e351.

Mauri, A., Strona, G., & San-Miguel-Ayanz, J. (2017). Eu-forest, a high-resolution tree occurrence dataset for europe. *Scientific data*, *4*, 160123.

Murphy, H. T., VanDerWal, J., & Lovett-Doust, J. (2006). Distribution of abundance across the range in eastern north american trees. *Global Ecology and Biogeography*, *15*, 63–71.

Nathan, R. & Muller-Landau, H. C. (2000). Spatial patterns of seed dispersal, their determinants and consequences for recruitment. *Trends in Ecology & Evolution*, *15*(7), 278–285.

Nathan, R., Safriel, U., & Imanuel, N. (2001). Field validation and sensitivity analysis of a mechanistic model for tree seed dispersal by wind. *Ecology*, *82*(2), 374–388.

Ovaskainen, O. & Meerson, B. (2010). Stochastic models of population extinction. *Trends in Ecology & Evolution*, *25*(11), 643–652.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Plummer, M. (2003). Jags: A program for analysis of bayesian graphical models using gibbs sampling.

Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology Letters*, *3*(4), 349–361.

Purves, D. W., Zavala, M. A., Ogle, K., Prieto, F., & Benayas, J. M. R. (2007). Environmental heterogeneity, bird-mediated directed dispersal, and oak woodland dynamics in mediterranean spain. *Ecological Monographs*, *77*(1), 77–97.

Reich, P. B. & Oleksyn, J. (2008). Climate warming will reduce growth and survival of scots pine except in the far north. *Ecology Letters*, *11*(6), 588–597.

Ruiz-Benito, P., Ratcliffe, S., Zavala, M. A., Martínez-Vilalta, J., Vilà-Cabrera, A., Lloret, F., ... Kändler, G., et al. (2017). Climate-and successional-related changes in functional composition of european forests are strongly driven by tree mortality. *Global change biology*, *23*(10), 4162–4176.

Sagarin, R. D., Gaines, S. D., & Gaylord, B. (2006). Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology & Evolution*, *21*(9), 524–530.

San-Miguel-Ayanz, J., de Rigo, D., Caudullo, G., Houston Durrant, T., & Mauri, A. (Eds.). (2016). *European atlas of forest tree species*. Luxembourg: Publication office of the European Union.

Schurr, F. M., Midgley, G. F., Rebelo, A. G., Reeves, G., Poschlod, P., & Higgins, S. I. (2007). Colonization and persistence ability explain the extent to which plant species fill their potential range. *Global Ecology and Biogeography*, *16*(4), 449–459.

Schurr, F. M., Pagel, J., Cabral, J. S., Groeneveld, J., Bykova, O., O'Hara, R. B., ... Zimmermann, N. E. (2012). How to understand species' niches and range dynamics: A demographic research agenda for biogeography. *Journal of Biogeography*, *39*(12), 2146–2162.

Snyder, R. E. & Chesson, P. (2004). How the spatial scales of dispersal, competition, and environmental heterogeneity interact to affect coexistence. *The American Naturalist*, *164*(5), 633–650.

Soberón, J. (2007). Grinnellian and eltonian niches and geographic distributions of species. *Ecology Letters*, *10*(12), 1115–1123.

Svenning, J.-C., Gravel, D., Holt, R. D., Schurr, F. M., Thuiller, W., Münkemüller, T., ... Normand, S. (2014). The influence of interspecific interactions on species range expansion rates. *Ecography*, *37*(12), 1198–1209.

Svenning, J.-C., Normand, S., & Kageyama, M. (2008). Glacial refugia of temperate trees in europe: Insights from species distribution modelling. *Journal of Ecology*, *96*(6), 1117–1127.

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

Talluto, M. V., Boulangeat, I., Vissault, S., Thuiller, W., & Gravel, D. (2017). Extinction debt and colonization credit delay range shifts of eastern North American trees. *Nature Ecology & Evolution*, *1*, 0182.

Thuiller, W., Lafourcade, B., Engler, R., & Araújo, M. B. (2009). Biomod–a platform for ensemble forecasting of species distributions. *Ecography*, *32*(3), 369–373.

bioRxiv preprint doi: https://doi.org/10.1101/748483; this version posted August 29, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Thuiller, W., Lavorel, S., & Araújo, M. B. (2005). Niche properties and geographical extent as predictors of species sensitivity to climate change. *Global Ecology and Biogeography*, *14*(4), 347–357.

Thuiller, W., Münkemüller, T., Schiffers, K. H., Georges, D., Dullinger, S., Eckhart, V. M., ... Schurr, F. M. (2014). Does probability of occurrence relate to population dynamics? *Ecography*, *37*(12), 1155–1166.

Trabucco, A. & Zomer, R. (2010). Global soil water balance geospatial database. *CGIAR Consortium for Spatial Information. Published online, available from the CGIAR-CSI GeoPortal at: http://www. cgiar-csi. org.*