Shaping the niche of *Taxus baccata*, a modelling exercise using biologically meaningful information

Pablo Sanchez-Martinez[1,2,*], Arnald Marcer[1,2], Maria Mayol[1,2] and Miquel Riba[1,2]

[1] CREAF, Cerdanyola del Vallès 08193, Spain

[2] Univ Autònoma Barcelona, Cerdanyola del Vallès 08193, Spain

*Corresponding author: p.sanchez@creaf.uab.cat

Abstract

Widely used Correlative Species Distribution Models (C-SDMs) usually make some simplifying assumptions, often failing to consider important ecological and evolutionary attributes potentially hindering the characterization of the species niche. Here, we use the tree species *Taxus baccata* to explore the effects of including biologically meaningful information on processes and features beyond purely abiotic factors that are expected to determine its niche and range size. To elucidate how these often neglected factors affect C-SDM results, we modelled the current niche in the species' southernmost European range using Maxent. More specifically, we included available basic information regarding biotic interactions, local adaptation and non-equilibrium demographic dynamics. The potential effect of biological interactions was introduced using habitat suitability of co-occurring tree species as predictive variables. Local adaptation was included modelling two distinct regional adaptive groups. We also used individual growth estimated under field conditions as a surrogate for demographic behaviour to control for the quality of model predictions and empirically assess the effect of biotic interactions. Including information on co-occurring tree species improved model performance and decreased the projected range size in most cases. These effects were not a result of biological interactions *per se*, but instead a consequence of co-occurring species accounting for fine-scale environmental variability not described by any of the climatic variables used. Considering local adaptation allowed detecting the role of different climatic variables in shaping the niche of each adaptive group that could potentially also act as selective pressures in the near future. Finally, and more importantly, we found that including populations that are probably currently found under non-equilibrium suboptimal conditions might largely overestimate the species niche.

39   **Keywords**: biotic interactions, local adaptation, plant performance, range size, Species

40   Distribution Models (SDM), *Taxus baccata*.

41

1. Introduction

About half a century ago, G. Evelyn Hutchinson (1957) provided a quantitative formalisation of the species ecological niche concept as the n-dimensional hyper-volume defined by abiotic and biotic variables where a species can potentially survive. From a demographic perspective, the Hutchinsonian niche of a species represents those habitats where its intrinsic population growth rate is zero or positive (Holt, 2009). Hutchinson's ideas were first related to Correlative Species Distribution Models (C-SDMs) in an early study assessing the climatic requirements of tree species (Booth, 1988), and this theoretical and methodological framework has since then been widely used. However, C-SDMs aim to describe patterns rather than mechanisms leading to the association between species occurrences and environmental data, making their ecological meaning hard to interpret (Singer et al., 2016).

C-SDMs have become the standard approach to predict species range dynamics due to the urgency of estimating ecological responses to rapid environmental change and the general lack of detailed individual-based information necessary to parameterize mechanistic models, which include explicit processes aiming at defining causality (Singer et al., 2016). Since C-SDMs are not directly linked to ecological and evolutionary processes, some of their assumptions may not always be fulfilled. Among the main assumptions that might not always hold are the following: (1) species' distributions are mainly shaped by climatic conditions; (2) the species niche is conserved over space and time; and (3) species' distributions will frequently be at equilibrium with the environment (Araújo and Peterson, 2012). However, biologically informed correlative models might be able to handle some of these limitations by considering basic information on certain potential ecological and evolutionary processes (Thuiller et al., 2014; Ehrlén and Morris, 2015). These approaches might be particularly useful for species where C-SDM

67   assumptions might not hold up and process-based information is rather limited or hard to

68   obtain. In these cases, the consideration of any available information about the species

69   biology when implementing C-SDMs may allow not only to test for the consequences of

70   potential deviations from model assumptions, but to provide meaningful ecological

71   insights (Dullinger et al., 2012; Giannini et al., 2012; Marcer et al., 2016; Talluto et al.,

72   2016).

73   Biotic interactions can sometimes override or modify the effects of climate

74   (Araújo and Luoto, 2007). For instance, weak competitors can be excluded by dominant

75   species from their optimal environmental conditions, while they might persist in more

76   extreme environments that the dominant competitors cannot occupy (Thuiller et al.,

77   2014). Although from a theoretical and empirical point of view biotic interactions are

78   essential in defining the species' niche (Wisz et al., 2013), their role in C-SDMs for

79   predicting species distributions is unclear, and probably depends on the spatial scale

80   (Pearson and Dawson, 2003) and the type of biotic interaction considered (Araújo and

81   Rozenfeld, 2014). Different methodological and statistical approaches have been

82   suggested to better assess the effect of biotic interactions (reviewed in Wisz et al., 2013),

83   but it remains rather challenging to infer their role in defining the niche, mainly because

84   of the confounding effects of missing abiotic variables and the temporal and spatial

85   dependency of interactions on environmental factors (Dormann et al., 2018). This might

86   be particularly problematic when the role of biotic factors changes according to the life-

87   history stage considered. Although C-SDMs that include co-occurrence information at

88   large spatial scales do not necessarily describe biotic interactions (Giannini et al., 2012),

89   they can help to put their potential importance in shaping species niche into perspective

90   (Dormann et al., 2018). Moreover, including information regarding the distribution of

91   potential biotic interactions increases the predictive power of the model, as it might

92   account for environmental axes not described by widely used climatic data (Zimmermann

93   et al., 2010; Giannini et al., 2012; Wisz et al., 2013).

94   Even though in commercial forestry the importance of provenance selection trials

95   has long been acknowledged, another common assumption behind most studies using C-

96   SDMs for predicting species ranges is that of niche conservatism, meaning that the

97   suitable niche space does not change over space and time. However, in many cases broad

98   climatic tolerances at the species level are generally comprised of narrower,

99   geographically-adapted tolerances at the scale of populations as a consequence of

100   differential selection pressures and/or population isolation (Wiens et al., 2009; Benito-

101   Garzón et al., 2011, 2019; Peterson et al., 2019). Some approaches have attempted to

102   incorporate intraspecific variation as a proxy for local or regional adaptations into species

103   niche modelling (reviewed in Peterson et al., 2019) by, for instance, modelling separately

104   different taxonomic units (Meynard et al., 2017) or genetic lineages (Marcer et al., 2016),

105   or even including phenotypic plasticity and adaptation over large geographical scales

106   (e.g., Valladares et al., 2014; Fréjaville et al., 2020). Including intraspecific variation can

107   account for deviations from the niche conservatism assumption, being especially useful

108   when modelling species where local adaptation is known to happen. In such cases,

109   modelling without taking into account this relevant evolutionary information could lead

110   to erroneous predictions or forecasts of future range dynamics, increasing (e.g., Pearman

111   et al., 2010; Benito-Garzon et al., 2011; Bush et al., 2016; Razgour et al., 2019) or

112   reducing (e.g., Atkins and Travis, 2010; Valladares et al., 2014) the projected

113   vulnerability to climate change.

114   Finally, C-SDMs assume that species are at equilibrium with current climatic

115   conditions, so that the species' range is expected to fill the entire suitable habitat

116   available. However, under non-equilibrium situations, species may be absent from

117   suitable environments (Booth et al., 2014). In fact, commercial forestry trials have

118   frequently shown that many tree species can grow successfully under climatic conditions

119   other that those found in their natural range (Booth, 2017), so that C-SDMs may thus

120   underestimate a specie's fundamental niche and its climatic adaptability. In other

121   situations, species may be present in environments that once were suitable but are

122   currently unsuitable (Dullinger et al., 2012). This last scenario might be particularly

123   important in long-lived organisms with declining or remnant population dynamics

124   (Eriksson and Eriksson, 2000), as well as at the "trailing" edge of species that are

125   currently contracting their ranges (Elith et al., 2010). In such situations C-SDMs might

126   overestimate the extension of the species niche, as some populations are wrongly assumed

127   to be under suitable conditions (Schurr et al., 2012). From a theoretical viewpoint, the

128   probability of occurrence or suitability projected by C-SDMs is expected to be related to

129   the species' demographic performance, measured as the intrinsic population growth rate

130   or some of their components (survival, growth or reproduction). Few studies have

131   incorporated functional or demographic traits and/or compared the relation between

132   demographic parameters and occurrence probability derived from C-SDM models

133   (Benito-Garzón et al., 2013). Nevertheless, some studies have reported that the

134   probability of occurrence or suitability does not always co-vary with trait expression as a

135   proxy of species performance, or with demographic parameters (Thuiller et al., 2009;

136   Dolos et al., 2015), as it would be expected if model assumptions were met.

137   In this study, we used the English yew (*Taxus baccata* L., Taxaceae) to assess the

138   effects of integrating biological meaningful features that are expected to determine the

139   species' distribution beyond purely abiotic factors on the habitat suitability derived from

140   C-SDMs. This species is an example of a non-model and non-commercial long-lived tree

141   with remnant and declining population dynamics, particularly in its southern range, for

142   which some relevant biological information is available. To that end, we evaluated the

143   effects of incorporating biological information into C-SDMs regarding (1) co-occurring

144   species niche distributions and field observations of inter and intraspecific competition,

145   (2) regional patterns of adaptation and (3) individual performance (growth). In particular,

146   we are addressing the following questions: (1) Does the inclusion of co-occurring species

147   niche suitability provide meaningful information about species interactions and/or

148   increase the predictive power of the model? (2) Does the inclusion of local adaptation

149   lead to different niche predictions as compared to the whole species approach? (3) Is the

150   consideration of individual performance useful to inform distribution models and assess

151   the predictive quality of C-SDMs? Our aim is not to provide an accurate description of

152   the species' niche, but to show that using basic and relevant biological information in C-

153   SDMs helps unravelling the importance of the ecological and evolutionary processes

154   considered and, more importantly, that it can also be a valuable management tool for

155   benchmarking the quality of much simpler predictive models.

156

157   **2. Materials and methods**

158   *2.1 Study species and area*

159   English yew (*T. baccata* L.) is a dioecious, slow-growing and long-lived gymnosperm

160   natural of Eurasian temperate and Mediterranean forests (Thomas and Polwart, 2003).

161   Despite its wide distribution, it forms small stands and/or isolated populations in many

162   parts of Europe, particularly in the Mediterranean area. In the Iberian Peninsula, the

163   species is found under quite variable environmental conditions across most of its

164   latitudinal range. In the south, where climate is characterized by higher temperature and

165   less precipitation compared to the north, populations are often small, sometimes with no

166   more than 20-150 reproducing trees that are located in north-facing slopes and shady

167   ravines or close to streams and mountain springs. In these populations regeneration is

168   almost absent, since seedling emergence and recruitment in English yew is partly limited

169   by water availability (Sanz et al., 2009), suggesting a high extinction risk in the coming

170   future. In contrast, the species can form relatively large and continuous stands in northern

171   latitudes, of up to 1,000-3,000 individuals, although it is often found growing with other

172   tree species in mixed forests of beech, pines and oaks. Potential biotic interactions with

173   these accompanying species may include facilitation for recruitment and growth during

174   the first sapling stages (García et al., 2000; García and Obeso, 2003), but also competition

175   for light under excessive canopy closure that can strongly limit sapling growth and

176   reproduction of adults (Svenning and Magård, 1999; Iszkuło, 2010; Iszkuło et al., 2012).

177   As in many other parts of its current range, gene flow among populations of the

178   Iberian Peninsula is limited, and neutral genetic diversity is highly structured both at the

179   local (Dubreuil et al., 2010) and the regional scale (González-Martínez et al., 2010;

180   Maroso et al., 2021). Within the Iberian Peninsula, adaptive variation in response to local

181   environmental drivers has been reported for some of the genes involved in taxol

182   biosynthesis (Burgarella et al., 2012) and, more recently, for phenological and growth

183   traits (Mayol et al., 2020).

184

185   *2.2 Individual performance in natural populations*

186   We assessed the performance of individuals in natural populations using tree growth,

187   estimated from increment cores for a total of 235 trees sampled in 25 natural populations

188   across the study area (Fig. 1a, Table 1). Most of these populations were visited to assess

189   demographic (population size and sex ratios) and genetic variability. Since many

190   populations are protected and the number of available trees per population is usually

191   limited, the number of trees sampled per population was rather variable (1-26 trees per

population; mean= 9.4; median = 9.0). Using a 5-mm increment borer, we usually extracted only one core at breast height in the upslope side of the tree. Increment cores were mounted on wooden supports, air-dried, sanded, and digitally scanned. Tree rings were identified using CooRecorder/CDendro software package (Larsson, 2013). Because of the limited number of sampled tress in many of the populations, we did not perform cross-dating of ring width variation. For all but two of the 235 sampled trees we also measured their diameter at breast height and categorized the tree as either uncovered (UC), i.e., dominant and codominant trees receiving almost full light from above ($< 25\%$ crown cover), or covered (intermediate or suppressed tree by another tree: $\geq 25\%$ crown cover). We also recorded the identity of the nearest neighbor dominant tree species. From this information, covered yew trees were further classified according to leaf size and leaf habit of the nearest neighbor: (i) needle-like perennials (NP, such as *Pinus spp.*), (ii) broad-lived perennials (BP, such as *Quercus ilex*), (iii) broad-leaved deciduous (BD, e.g. *Fagus sylvatica*), and (iv) conspecifics (T, *T.baccata*).

Information on radial growth increments was converted into basal area increments during the last five years (BAI5, in cm^2). We then used a general linear mixed model to assess the potential effect of the type of crown cover on BAI5. The model included cover type (UC, NP, BP, BD and T), individual size (BA: estimated 5-year's previous basal area in cm^2) and their interaction as covariates, and population identity as a random factor. Both BAI5 and BA were ln-transformed. The analysis was performed with the *lmer* function of the *lme4* package version 1.1 (Bates et al., 2015) in R (R Core Team, 2019). Tests for fixed effects were obtained using the *lmerTest* package version 1.1 (Kuznetsova et al., 2016) with the Kenward-Roger's approximation for denominator degrees of freedom for the F statistics (Kenward and Roger, 1997).

216   We also assessed the potential association between plant performance and

217 climate. We used stepwise regression *(StepReg* R package version 1.1, Li et al., 2020)

218 between mean BAI5 per population and average precipitation and daily mean minimum

219 and maximum temperatures during spring (March-April), summer (June-August),

220 autumn (September-November) and winter (December-February).

221

222 *2.3 Occurrence data*

223 Occurrence records of *T. baccata* in the Iberian Peninsula were obtained from individual

224 GPS data points and 1-km spatial resolution records. Individual GPS points included data

225 collected by the authors from 2005 to 2015, the LIFE-BACCATA project (LIFE15

226 NAT/ES/000790), and several local botanical experts, as well as information available in

227 the databases of the Valsaín Clonal Bank (Spanish Ministry of Agriculture, Fisheries and

228 Food) and the General Directorate for the Natural Environment (Valencian Autonomous

229 Community). One-km spatial resolution records were available or reported in Serra and

230 García (2012), ANTHOS (Plant Biodiversity Data bank of Spain, http://www.anthos.es/)

231 and BIOCAT (Biodiversity Data Bank of Catalonia, http://biodiver.bio.ub.es/biocat/)

232 databases. To standardize occurrence data, individual GPS data points were coarsened to

233 a 1-km spatial grid, finally obtaining 1,817 unique 1-km resolution points from the

234 original 9,017 occurrence GPS records (Fig. 1a).

235

236 *2.4 Climatic and biotic C-SDM predictors*

237 We considered two types of predictive variables: those related to climatic conditions and

238 those related to biotic interactions (biotic variables hereafter). Two types of datasets

239 concerning predictor variables were prepared: a) a set with only climatic variables, *Clim*,

240 and b) a set with the same climatic variables and biotic variables, *ClimBio*. Climatic

241 variables were downloaded from the Digital Climatic Atlas of the Iberian Peninsula

242 (Ninyerola et al., 2005) in February 2018. We used the mean value to aggregate the

243 original spatial resolution of 200 m to 1,000 m, and then generated the Bioclim set of

244 variables (Booth et al., 2014) using the *bioclim* function of the R package *dismo* version

245 1.1 (Hijmans et al., 2017). We chose a set of five bioclimatic variables with known or

246 alleged ecological relevance for the species (see Svenning and Skov, 2004; Sanz et al.,

247 2009; Cedro and Cedro, 2015; Mayol et al., 2020): annual precipitation, summer

248 precipitation, precipitation seasonality, mean winter temperature and temperature

249 seasonality.

250 As for biotic variables, we used existing niche model projections (habitat

251 suitability) of co-occurring tree species (Atlas of Topo-Climatic Suitability of Woody-

252 Plants; Ninyerola et al., 2010, accessed in April 2018). We selected a set of tree species

253 known to co-occur with *T. baccata* throughout the Iberian Peninsula: *Fagus sylvatica,*

254 *Quercus humilis, Quercus pyrenaica, Pinus pinaster* and *Pinus nigra*. All the biotic and

255 abiotic predictive variables selected showed low collinearity (all pair-wise Pearson

256 correlation coefficients: $r < 0.7$) for the combined set of occurrences and background

257 points.

258

259 *2.5 Incorporating geographical patterns of adaptation*

260 Mayol et al. (2020) suggested the existence of local/regional adaptation of *T. baccata* to

261 climatic conditions in the Iberian Peninsula. In particular, plant growth and reproduction

262 (male strobili maturation) recorded in a common environment were found to be related to

263 several temperature variables. Their results suggested that populations in the Iberian

264 Peninsula form an adaptive cline from *Continental* populations adapted to colder

265 temperatures to *Mild* populations adapted to warmer temperatures. Since variation in

266   shoot growth (often a good fitness proxy for forest trees) among geographical

267   provenances was mainly associated with temperatures experienced during the colder

268   season, we used the mean winter (December, January, February) temperature to classify

269   the experimental provenances from Mayol et al. (2020) into either *Continental* or *Mild*.

270   We used the cut-off value of 3.5ºC to obtain the same frequency of provenances in each

271   class. This same classification criterion was then applied to the dataset of natural

272   occurrences: 833 below 3.5ºC were classified as *Continental* and 984 above or equal to

273   3.5ºC as *Mild* (Fig. 1a). To assess the effects of including adaptive variability on model

274   predictions we used both sets of adaptive groups to generate separate C-SDMs models

275   and compared them with models using the whole set of occurrences.

276

277   *2.6 Using individual performance to assess predictive quality of C-SDMs*

278   Based on the relationship between individual performance and climate (see Results), we

279   fitted a regression model including spring precipitation and maximum autumn

280   temperature as predictors and growth (BAI5) as the response variable. We then used this

281   model to predict the growth at each of the 1,817 natural occurrences. To assess the quality

282   of the suitability predictions arising from C-SDMs, and to determine the set of

283   occurrences that maximize the relationship between occurrence-derived suitability and

284   growth-based estimates of habitat quality, we created nine (filtered) datasets including

285   the 90%, 80%, 70% ... 10% of the top predicted growth values (BAI5) (Figs. 1b,c,d and

286   Fig. S1).

287

288   *2.7 Correlative species distribution modelling and range size*

289   A scheme of the whole modelling approach is provided in Fig. 2. We extracted values at

290   each of the 1,817 locations for all predictors and created different datasets and models by

291   faceting our data along the following criteria: 1) the type of predictors (*Clim* and *ClimBio*

292   datasets), 2) the incorporation of geographical patterns of adaptation (*Continental* and

293   *Mild* adaptive groups, as well as considering the species as a whole, SP), and 3) the use

294   of all occurrences (unfiltered models: datasets with 100% of the occurrences) and taking

295   into account individual performance with 10% filtering steps according to predicted

296   growth (BAI5), as described above. These combinations resulted in 60 datasets, which

297   were used as input to the corresponding 60 distribution models: two types of predictors

298   (*Clim* and *ClimBio*) x three groups (*Continental*, *Mild* and SP) x ten occurrence datasets.

299   We generated a total of 20 additional distribution model maps that were not a direct

300   product of the modelling algorithm but *a posteriori* composition of the species

301   distribution after merging the outputs of the distribution models for the *Continental* and

302   *Mild* adaptive groups (C+M). These models were generated by keeping the highest value

303   of the two models in each cell as in Marcer et al. (2016).

304   We used Maxent, version 3.3.3k, with the *maxent* function in the *dismo* R package

305   version 1.1 (Hijmans et al., 2017) to build our distribution models. Maxent is a presence-

306   background modelling algorithm based on the maximum entropy principle (Phillips et al.,

307   2006; Elith et al., 2011) which has been shown to outperform other correlative algorithms

308   of its type (Elith et al., 2006). We used Maxent with its default settings except for feature

309   types, for which only the hinge type was selected; hence similar to a general additive

310   model (GAM) (Elith et al., 2011). For each of the 60 modelling datasets, 30% of the

311   occurrences were randomly separated for a final evaluation of predictive performance and

312   the remaining 70% occurrences were used to train each model. We selected a single

313   random set of 10,000 points for which we extracted their environmental values and added

314   them as background to each modelling set of occurrences. We used Maxent's permutation

315   importance to assess the contribution of each variable to the model, i.e. the drop in

training AUC normalized as percentage when generating the model with its values randomly permuted. AUC is the receiver operating characteristic area under the curve, a threshold-independent measure of predictive performance. Marginal response curves of each variable were produced by plotting each predictor variable against predicted suitability while maintaining the rest of the variables at their average sampling value.

To assess the predictive performance of models, we calculated the test-AUC using the *evaluate* function in the *dismo* R package version 1.1 (Hijmans et al., 2017) on the 30% of the test points. In addition, we calculated the Spearman correlation coefficient between the suitability value obtained for each model and the predicted growth (BAI5) using the *cor.test* function (R Core Team, 2019). Then, we selected the best model for each combination of predictors (*Clim* and *ClimBio*), groups (*Continental*, *Mild*, SP) and the ten occurrences (one unfiltered and nine growth-filtered) datasets. We considered as best models those with the highest significant suitability-growth correlation and a test-AUC value equal or higher than 0.9, a value indicating very good predictive performance according to Swets (1988).

Finally, we estimated the range size for the unfiltered and the best models by assuming that the available range for the whole species or adaptive group was proportional to the predicted suitability of their models as in Oney et al. (2013). Range size ($Area_{tot}$) was calculated for each model by multiplying the predicted suitability values of each j grid cell by grid cell area (i.e., 1 km^2), and summing over the M analyzed cells:

$$Area_{tot} = \sum_{j=1}^{M} P_j * Area_j$$

3. Results

3.1 Individual performance in natural populations

Individual values for growth during the last five years (BAI5) ranged from 2.9 to 355.4 cm^2 (mean = 45.9 cm^2; median = 30.0 cm^2). Differences among individual trees in BAI5 were related to the size of the tree (β_{BA} = 0.33; p < 0.001) and the type of tree cover (p < 0.001). We did not find a significant interaction between tree size and type of cover (p = 0.12). As expected, the highest mean growth values were those of uncovered trees, whereas the lowest growth corresponded to individuals suppressed by conspecifics (Fig. 3). Individuals partially or fully covered by needle-like perennials showed similar growth values to uncovered trees, and those covered by broad-leaved perennial species showed similar mean growth values to conspecifics ones. Trees covered by broad-leaved deciduous species showed intermediate values.

Stepwise regression of mean population growth (BAI5) values on climatic predictors only selected two dependent variables: the first step included precipitation during spring, and the second step included mean daily maximum temperature during autumn. The estimated bivariate regression model (R^2 = 0.69) suggested a significant positive relationship between BAI5 and spring precipitation (partial regression coefficient = 0.453; t = 6.87; p <0.001) and a positive association with maximum autumn temperature (partial regression coefficient = 11.021; t = 4.22; p <0.001). Mean performance (BAI5) values per population where positively correlated with estimated census sizes (Spearman's Rho = 0.65; p <0.001) (see Table 1). Predicted BAI5 values for the 1,817 1-km grid species occurrences ranged between -41.979 and 247.251 cm^2, 29 of them with mean predicted values lower than zero and 154 not different from zero (95% CI). These occurrences corresponded to localities in the lowest range of spring precipitation: all lower than 300 mm and 75% of them lower than 200 mm.

362

363   *3.2 Model performance and selection of best models*

364   Model accuracy was high, both for models excluding (*Clim*) and including biotic

365   variables (*ClimBio*) (Table 2, Table S1): mean test-AUC ± SD values were 0.937 ± 0.042

366   and 0.950 ± 0.029 for *Clim* and *ClimBio* models, respectively (Table S1). In general,

367   higher accuracy was obtained when modelling the *Continental* and *Mild* adaptive groups

368   separately than considering the whole set of occurrences (SP). For the *Continental* group,

369   test-AUC values were always ≥ 0.970 (mean 0.982 ± 0.007), while they ranged from

370   0.886-0.995 for the *Mild* group (mean 0.939 ± 0.031), and from 0.874-0.989 for the SP

371   models (mean 0.924 ± 0.036; Table 2, Table S1). Unfiltered models, i.e. those based on

372   datasets with all the occurrences, had lower predictive power (test-AUC range from 0.880

373   to 0.973, mean 0.917 ± 0.037) than those built on growth-filtered datasets (test-AUC

374   range from 0.874 to 0.995, mean 0.947 ± 0.036). A particularly good accuracy was

375   obtained for those filtered models built from datasets with 30% or less of the best

376   predicted growth values (test-AUC range from 0.953 to 0.995, mean 0.977 ± 0.014; Table

377   S1).

378   When compared with filtered datasets, none of the unfiltered models (100%

379   occurrences) showed the highest significant suitability-growth correlation (Table 2).

380   Furthermore, for most unfiltered models the correlation between predicted suitability and

381   growth was negative, with the exception of those built with the *Continental* group of

382   occurrences, for which the Spearman's rank correlation coefficient was rather low (Table

383   2). Within *Clim* models, the highest significant suitability-growth correlation for the

384   *Continental* adaptive group was found for the filtered model with the 10% of the top

385   predicted growth values (Spearman's Rho = 0.811, p <0.001). The best models for *Mild*

386   and SP groups were those based on the 30% (Spearman's Rho = 0.838, p <0.001) and

17

387   20% (Spearman's Rho = 0.742, p <0.001) of the top predicted growth values, respectively

388   (Table 2). For C+M models, the highest significant suitability-growth correlation was

389   found for the model with 40% of the top predicted growth values (Spearman's Rho =

390   0.695, p <0.001). As for the *ClimBio* predictor models, the best ones for the *Continental*,

391   *Mild* and SP groups were found when using the 30% (Spearman's Rho = 0.805, p <0.001),

392   30% (Spearman's Rho = 0.846, p <0.001) and 20% (Spearman's Rho = 0.792, p <0.001)

393   of the top growth values, respectively (Table 2). Finally, the highest significant

394   suitability-growth correlation for C+M models was found for the model with 20% of the

395   top predicted growth values (Spearman's Rho = 0.774, p <0.001).

396

397   *3.3 Response to predictor variables*

398   The variables with higher permutation importance for *Clim* models were mean winter

399   temperature and annual precipitation (Table 3, Table S2). Mean winter temperature was

400   particularly relevant in all the models for the *Continental* adaptive group (85.9% and

401   64.3% for the unfiltered and the best model, respectively), positively influencing

402   predicted suitability (Table 3, Fig S2). Annual precipitation had the largest and positive

403   importance in SP models (53.7% and 76.9% for unfiltered and the best model,

404   respectively) (Table 3, Fig. S2). Annual precipitation also showed a high and positive

405   influence in the models for the *Mild* adaptive group (38.4% and 67.2% for unfiltered and

406   the best model, respectively), although in this case mean winter temperature showed an

407   important and negative influence on predicted suitability in the model using the unfiltered

408   dataset (27.9%) (Table 3, Fig. S2).

409   As for the models build including additional biotic predictors (*ClimBio*), the sign

410   and permutation importance of annual precipitation and mean winter temperature were

411   quite similar to those described above for climatic-only predictors, except for the

412   unfiltered *Mild* and SP occurrence datasets (Table 3). In this case, and for the *Mild*

413   adaptive group, the presence of *Pinus pinaster* was the predictor variable with the largest

414   permutation importance (24.7%), negatively affecting suitability (Fig. S3), followed by

415   the negative effect of mean winter temperature (18.0%) and the positive effect of annual

416   precipitation (17.9%). For SP group of occurrences, the two variables with the largest

417   relative importance were the presence of *Fagus sylvatica* (46.1%) and *Quercus humilis*

418   (15.0%), showing a positive and negative influence, respectively, on the predicted

419   suitability of *T. baccata* (Table 3. Fig. S3).

420

421   *3.4 Predicted suitability and range size*

422   Among all the models using unfiltered data, those built with both climatic and biotic

423   variables (*ClimBio*) always predicted lower suitability for *T. baccata* than climatic-only

424   (*Clim*) models (Fig. 4), resulting into range size reductions around 30-40% (Fig. 5a). The

425   predicted suitability and range size were lower for the *Continental* adaptive group

426   compared to the *Mild* one (Figs. 4 and 5a). Habitat suitability for the former was mostly

427   restricted to the main mountain ranges present in the Iberian Peninsula, characterized by

428   high elevations and/or continental cold climates, while suitable areas for the latter were

429   found at lower elevation mountain ranges located near the Mediterranean and Cantabrian

430   coasts, characterized by warmer climates (Fig. 4). Due to the little overlap between the

431   predicted suitability for each group, the predicted range size for C+M models was close

432   to the sum of range sizes obtained for each group separately (Fig. 5a). The predicted range

433   size obtained for the species as a whole (SP) did not differ substantially from that resulting

434   from C+M models (Fig. 5a), although it was slightly smaller for *Clim* models (Fig. 5a).

435   Overall, and compared with unfiltered models, the predicted suitability and range

436   size of the best filtered models was strongly reduced (Figs. 4 and 5b). Projected range

437 size decreased between 52% to 81%, depending on the type of predictive variables used,

438 whether the presence of adaptive groups was considered or not, and the type of adaptive

439 group. In general, the reduction was more or less proportional to the percentage of

440 occurrences used to train the models.

441

442 *3.5 Predicted suitability and individual performance*

443 The correlation between predicted suitability and BAI5 improved progressively when

444 filtering (Fig. 6, Table S3). For the *Continental* adaptive group, the maximum correlation

445 was reached when models were built with the 10-30% of the top predicted growth values,

446 while for the *Mild* adaptive group, SP and composite C+M datasets the correlation slightly

447 decreased when the strongest restrictive filtering categories were applied (Fig. 6).

448 Spearman's rank correlation coefficients increased from low values close to zero to

449 positive values around 0.7-0.8, depending on the model (Fig. 6, Table S3). The effect of

450 filtering was particularly pronounced for *Mild* and SP groups of occurrences, changing

451 from slightly negative to highly positive correlations between predicted suitability and

452 growth (Fig. 6). The patterns of changes in correlations associated to filtering were very

453 similar for both *Clim* and *ClimBio* type of model predictors (Fig. 6).

454

455 **4. Discussion**

456 Purely data-driven models based exclusively on occurrence and climatic data are

457 nowadays routinely employed for both basic and applied biodiversity research, including

458 the detection of niche shifts, conservation planning and reserve design, land use planning

459 and restoration, and to predict the impact of anthropogenic climate changes (see Elith and

460 Franklin, 2013; Van Echelpoel et al., 2015). Since SDMs are conceptually underpinned

461 in basic ecological concepts and assumptions (Elith and Franklin, 2013), failure to

462   recognise the linkage between these assumptions and common modelling techniques

463   might be a critical issue limiting their predictive power and overall utility. In that sense,

464   data-driven models, without substantial integration of existing ecological knowledge,

465   have long been criticized on several grounds, among them those associated with the role

466   of interspecific interactions, dispersal and ecological equilibria, time lags, evolution and

467   sampling of niche space (e.g., Sinclair et al., 2010). By combining data with basic

468   ecological and evolutionary knowledge models can become more process-based, thereby

469   supporting predictions with a more sound basis (Van Echelpoel et al., 2015).

470   Unfortunately, novel and more refined methodological approaches accounting for key

471   ecological and evolutionary processes (e.g., Benito-Garzón et al., 2019) might only be

472   applied to species for which extensive empirical knowledge is or might be developed.

473   However, biologically-informed correlative approaches might still be a useful alternative

474   when only limited information is available.

475   Here, using a newly assembled database for *T. baccata* in the Iberian Peninsula,

476   we take advantage of the ecological and evolutionary information available on the species

477   to gauge the predictive power of C-SDMs when some of the basic assumptions are not

478   met. In doing so In doing so, we offer a guiding protocol that can be further explored and

479   developed in some other similar species, in particular long-lived tree species that might

480   be currently experiencing non-equilibrium demographic dynamics in a large part of their

481   range. Although the current approach is rather demanding, its applicability will depend

482   on the quantity and quantity of meaningful data available. We show that modifying some

483   of the most basic assumptions underlying C-SDMs, i.e. niche conservatism, equilibrium

484   dynamics, and that the niche is mainly constrained by a few and simple climatic variables,

485   had varying effects on the predicted suitability and geographic species' range, as well as

486   on the role of the predictive climatic variables included to define the current niche. Below

487 we discuss the potential importance, meaning and implications of addressing each of the

488 pitfalls or hidden assumptions according to the results obtained.

489

490 *4.1 The role of climatic and biotic predictors*

491 As expected based on several observational and empirical studies (Sanz et al., 2009;

492 Cedro and Cedro, 2015), including the significant relationship between mean population

493 growth (BAI) and spring precipitation found in our study, the results of the C-SDMs

494 showed that the presence of *T. baccata* was positively associated with annual rainfall. In

495 some models, winter temperature was also found be a relevant climatic variable, with

496 either positive or negative effects depending on the model and adaptive group considered.

497 Including biotic variables as predictors had variable effects on the relative importance of

498 climatic predictive variables in explaining species occurrence, as well as on the resulting

499 niche projection, depending also on the consideration of adaptive groups (*Continental*,

500 *Mild*, and SP).

501 When occurrences were not filtered according to predicted individual

502 performance (growth), including predictive biotic variables did not change the overall

503 positive or negative effect of the climatic variables, but they decreased both their relative

504 importance and the total projected range size. This suggests that the use of the standard

505 correlative model based on a few climatic predictive variables might not be fully

506 appropriate and that some relevant information is missing when defining the species'

507 niche. For example, when modelling the niche without taking into account adaptive

508 groups (SP) and using only climatic variables (*Clim*), annual precipitation was the most

509 important variable, positively associated with the probability of occurrence. However, its

510 importance decreased when biotic variables (*ClimBio*) were included, and *F. sylvatica*

511 suitability became the best predictor, showing also a positive effect on species occurrence.

512   These results did not agree with our empirical field data indicating that the presence of

513   broad-leaved species such as *F. sylvatica* had a negative effect on *T. baccata* growth (BD

514   in Figure 3), a pattern that has been previously reported in other studies (e.g., Piovesan et

515   al., 2009). Therefore, the observed positive association suggests that, instead of

516   accounting for the direct effects of a true biotic interaction, the predicted suitability of *F.*

517   *sylvativa* might act as an informative climatic proxy, whose effects might occur at a much

518   finer spatial scale. We cannot completely rule out, however, real biotic positive

519   interactions between both species. Even though *F. sylvatica* is expected to have a negative

520   effect on *T. baccata* through competition for resource acquisition, it might also positively

521   interact with *T. baccata* during the early stages of the life cycle trough "nursering" effects

522   on recruitment and establishment (García and Obeso, 2003; Iszkulo, 2010). If this is true,

523   these species may coexist except in those places where broadleaves form dense

524   populations/canopies and competition for light may be severe (Svenning and Magård,

525   1999), definitely excluding *T. baccata* (Ruprecht et al., 2010).

526   The relative importance and effect of including biotic predictors when using all

527   the available occurrence data (unfiltered models) was also different when modelling

528   discrete adaptive groups (i.e., *Mild* or *Continental*). For the *Continental* group, none of

529   the biotic predictors was found to be strongly associated with the probability of

530   occurrence, and mean winter temperature was the most important climatic predictive

531   variable, regardless of the set of predictor variables employed. However, the inclusion of

532   biotic variables produced a reduction of 31% in the projected range size of this adaptive

533   group. In contrast, when modelling the *Mild* adaptive group, the relative importance of

534   climatic variables (annual precipitation and mean winter temperature) was much more

535   reduced when biotic predictors were included. In this case, the projected range size was

536   reduced by 40% and the predictor variable with the highest relative importance was the

projected distribution of *P. pinaster*, which had a strong and negative effect on the probability of occurrence of *T. baccata*. As in the case of *F. sylvatica*, this negative relationship was not supported by field observations, which suggested a non-significant effect of needle-like perennials on *T. baccata* growth (NP in Fig. 3). These results likely indicate that this predictor would also not be truly representing a biotic interaction, but a climate proxy representing warm, open and/or hot-dry conditions with sandy soils, conditions that are not suitable for *T. baccata* persistence and where *P. pinaster* may reach a high suitability.

Our findings support previous results (Araújo and Rozenfeld, 2014; Godsoe et al., 2017) suggesting that negative interactions such as competition are not easy to capture at relatively low-resolution scales. However, in some cases, including the suitability of co-occurring species provides new dimensions of the environmental space, gathering niche requirements that constrain niche distribution and are not reflected in a simpler climatic characterization at low-resolution scales. In any case, it is worth noting that other studies have reported the ability to capture biotic interactions at similar spatial scales, especially when positive biotic interactions, such as diet resource distribution (De Araújo et al., 2014) or pollination (Giannini et al., 2012), are involved. In this context, it is worth noting that although *T. baccata* has traits favouring zoochory (i.e., fleshy fruits), most genetic analyses suggests that effective dispersal and/or recruitment are currently rather limited in the Iberian Peninsula (Dubreuil et al., 2010; González-Martínez et al. 2010; Maroso et al., 2021). Several biotic processes have been implicated in explaining this pattern, among them the activity of avian frugivores (Labrave & Garcia, 2015), the presence of other fleshy-fruited plants (García et al. 2000), and herbivore pressure by livestock and wild ungulates (Piovesan et al., 2009). We therefore suggest that information on these

561 biotic processes is worth exploring and incorporating in future modelling studies when

562 available.

563

564 *4.2 Local adaptation*

565 Previous studies suggested the presence of local adaptation of *T. baccata* in the Iberian

566 Peninsula (Mayol et al., 2020). In particular, phenotypic patterns of variation in shoot

567 growth and phenological development measured in a common garden environment

568 provided evidence of two basic climatic adaptive groups, with populations from

569 continental environments growing faster and reproducing earlier than those from mild,

570 temperate, coastal regions (Mayol et al., 2020). Our modelling results suggest that, while

571 cold temperature is the main climatic pressure constraining the presence of the

572 *Continental* group, warmer temperatures and water availability, i. e. drought, might be

573 the most important factor limiting the distribution of the *Mild* group.

574 Dendroclimatological analyses performed in populations from continental climates in

575 central-eastern Europe have similarly indicated that the dominant factor affecting

576 individual performance in *T. baccata* are thermal conditions during winter, and that

577 summer drought is an additional secondary factor limiting growth in dry locations (Cedro

578 and Cedro, 2015, and references therein). Therefore, both water deficit and high winter

579 temperatures are expected to act as the most important climatic selective pressures in the

580 near future in populations of the *Mild* and *Continental* adaptive groups, respectively.

581 Importantly, modelling the species as a whole (SP), without considering intraspecific

582 adaptive variation, greatly reduced the potential effect of temperature on the distribution

583 of the species, and the requirements of the *Continental* group were masked by those of

584 the more widely distributed *Mild* group. This could produce misleading results when

585 projecting the current niche space into the future or other geographical locations, since

586   the species treated as a whole does not truly represent the species physiological

587   limitations. The impact of considering adaptive groups in the current extent of the species'

588   range and suitability is rather low, as compared to the other factors and predictive

589   variables considered. However, this intraspecific evolutionary variability may be crucial

590   to better evaluate the potential impact of climate change scenarios on the species' range

591   distribution (Pearman et al., 2010). Previous studies show that for some species, the

592   incorporation of local adaptation can increase (e.g., Bush et al., 2016) or decrease (e.g.

593   Valladares et al., 2014) suitability projections under future conditions. Then, even though

594   we did not explicitly explore the effects of intraspecific adaptive variability on niche

595   projections in future climate scenarios, they could be crucial to better assess conservation

596   practices.

597   However, we must acknowledge the limitations of our approach, since adaptation

598   to local conditions probably follows an adaptive cline (see Mayol et al., 2020), rather than

599   conforming to the two-group classification employed. Our aim was not to accurately take

600   into account the patterns or clines of adaptation as described in Mayol et al. (2020), but

601   to explore the potential effects of adaptation on niche characterization as compared to

602   much simpler C-SDM models that do not account for evolutionary process or spatial niche

603   differentiation. Future work using specific clines or other complex and continuous

604   geographical patterns of adaptation to assess their effects on *T. baccata*'s niche space

605   could be of great interest.

606

607   *4.3 Non-equilibrium demographic dynamics*

608   The relationship between species habitat suitability as measured by C-SDMs and

609   demographic performance has seldom been tested and, when tested, the results are not

610   usually conclusive (Thuiller et al., 2009; Thuiller et al., 2014; Dolos et al., 2015). We

611   explored the importance of using estimates of individual growth as a way to inform C-

612   SDMs with an inferred measure of habitat suitability. Our analyses suggest that, for *T.*

613   *baccata* in the Iberian Peninsula, the commonly used C-SDMs may not in general truly

614   reflect habitat suitability, contrary to what is expected when modelling the Hutchinsonian

615   niche (i.e., environmental space where the species can grow and persist; Hutchinson,

616   1957; Holt, 2009). We only found a significant, albeit rather low, positive relationship

617   between suitability and projected growth when modelling separately the habitat suitability

618   of the *Continental* adaptive group. For the rest of modelling groups considered (*Mild*, SP,

619   and C+M), the correlation was not significant or, in most cases, significantly negative.

620   This furthermore indicates that the positive correlation between habitat suitability and

621   growth for the *Continental* group could have been masked by the low correlation reported

622   for the more widely distributed *Mild* group when local adaptation was not considered

623   (i.e., species level results, SP). In fact, the predicted growth was null for some populations

624   of the *Mild* group in locations were spring precipitation was low, suggesting that some of

625   the occurrences used to train the C-SDMs may actually be outside of the species

626   Hutchinsonian niche.

627   Therefore, we show that the standard C-SDM approach does not successfully

628   incorporate functional or demographic attributes of *T. baccata* in the Iberian Peninsula, a

629   finding that agrees with some previous work on other species (Thuiller et al., 2009; Lloret

630   et al., 2013; Thuiller et al., 2014; Dolos et al., 2015; Swab et al., 2015). However, our

631   methodology, based on filtering different proportions of occurrences according to their

632   predicted performance, improved the relationship between growth and suitability,

633   resulting in a reduction of the range size to those places where the species might be under

634   more optimal conditions. This is particularly important for populations located in the

635   species trailing range edge, as in our study case, where filtering by growth had a large

636   impact in reducing the estimated range size, helping to discriminate non-suitable areas

637   from more suitable ones and providing complementary and useful information to that

638   reported by standard C-SDMs.

639   Nevertheless, the filtering method implemented has some important caveats that

640   have to be considered. First, instead of individual growth, other demographic parameters

641   more closely related to population persistence (e.g., survival, reproduction or a joint

642   estimate) would better capture which populations are likely to be outside the

643   Hutchinsonian niche. Second, the relationship between growth and climate was modelled

644   using a simple linear function, and a more realistic sigmoidal function would probably be

645   more appropriate. Third, our population and individual tree sampling for estimating

646   growth in natural populations might also be positively biased towards populations with a

647   higher number of individuals and/or individuals that are more accessible and easier to

648   measure, which usually happen to be the ones under more favorable conditions and higher

649   growth. It is certainly difficult to accurately estimate growth using growth rings for

650   individuals that are under harsh adverse conditions, as some growth rings might be

651   missing or be too small to be detected, leading to errors in assessing the time frame

652   considered (in our case the last five years). However, these limitations and biases would

653   result in an overestimation of growth for the given period, so we believe that our results

654   are rather conservative regarding the relationship between growth and suitability, and that

655   the true relationship might be even lower. In any case, improved data availability may

656   help to better handle these limitations and it might be worth rerunning these analyses as

657   soon as more data becomes available.

658

5. Conclusions

Including individual or population features (e.g., ecological, demographic, functional, etc.) into C-SDMs can be useful even when the data available is rather limited, helping to test its consistency and to detect potential deviations from model assumptions. This is of special interest for non-commercial and non-model species for which there is not high-quality data to build mechanistic models, but enough for informing C-SDMs. This could be particularly useful for species showing non-equilibrium dynamics that might be currently occupying habitats that are suboptimal for their persistence. By including some of this basic information into *T. baccata* niche modelling we obtained a wide range of outcomes, showing how processes that are usually not considered when modelling distributions using C-SDM might influence its niche distribution at its southern range. In particular, the inclusion of patterns of adaptation and, notably, individual performance, can have strong effects on *T. baccata*'s niche projection. Although competition between coexisting species could not be successfully included into C-SDMs, the incorporation of the niche distribution of co-occurring species was very useful in capturing undetected environmental niche axes, improving model performance and having an important effect on estimates of range size. We also report evidence that some occurrences in the study area may be under non-suitable conditions, leading to an overestimation of the overall species niche when used to train C-SDMs.

We stress the value of including meaningful biological information in the C-SDM framework in order to better inform conservation and management policies. Though not explicitly addressed in this study, we expect that failure to recognize the limitations of simple C-SDMs when the main assumptions do not hold can have a similarly strong impact on the quality of predictions under future scenarios of climate change. On the plus side, however, the suggested protocol might be highly useful for both evaluating the

684 quality of predictions and posing new lines of inquiry concerning the biology,

685 conservation and management of the species considered.

686

699

700 **CRediT authorship contribution statement**

701 **P. Sánchez-Martínez:** Formal analysis, Writing - original draft, Writing - review &

702 editing. **A. Marcer:** Formal analysis, Writing - review & editing. **M. Mayol:**

703 Conceptualization, Funding acquisition, Resources, Writing - original draft, Writing -

704 review & editing. **M. Riba:** Conceptualization, Formal analysis, Funding acquisition,

705 Resources, Supervision, Writing - original draft, Writing - review & editing.

706

References

707 Araújo, M.B., Luoto, M., 2007. The importance of biotic interactions for modelling

708 species distributions under climate change. Global Ecol. Biogeogr. 16, 743–753.

709 https://doi.org/10.1111/j.1466-8238.2007.00359.x

710

711 Araújo, M.B., Peterson, A.T., 2012. Uses and misuses of bioclimatic envelope modeling.

712 Ecology 93, 1527–1539. https://doi.org/10.1890/11-1930.1

713 Araújo, M. B., Rozenfeld, A.,2014. The geographic scaling of biotic interactions.

714 Ecography 37, 406–415. https://doi.org/10.1111/j.1600-0587.2013.00643.x

715 Atkins, K.E., Travis, J.M.J., 2010. Local adaptation and the evolution of species' ranges

716 under climate change. J. Theor. Biol. 266, 449–457.

717 https://doi.org/10.1016/j.jtbi.2010.07.014

718 Benito-Garzón, M., Ruiz-Benito, P., Zavala, M.A., 2013. Interspecific differences in tree

719 growth and mortality responses to environmental drivers determine potential species

720 distributional limits in Iberian forests. Global Ecol. Biogeogr. 22, 1141–1151.

721 https://doi.org/10.1111/geb.12075

722 Booth, T. H., 2017. Assessing species climatic requirements beyond the realized niche:

723 some lessons mainly from tree species distribution modelling. *Clim. Change* **145**, 259–

724 271. 10.1007/s10584-017-2107-9

725 Booth, T. H., Nix, H. A., Busby, J. R. & Hutchinson, M. F., 2014. Bioclim: The first

726 species distribution modelling package, its early applications and relevance to most

727 current MaxEnt studies. *Divers. Distrib.* **20**, 1–9. 10.1111/ddi.12144

728 Booth, T. H., Nix, H. A., Hutchinson, M. F. & Jovanic, T., 1988. Niche analysis and tree

729 species introduction. *For. Ecol. Manage.* **23**, 47–59. 10.1016/0378-1127(88)90013-8

730 Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting Linear Mixed-Effects

731 Models Using lme4. J. Stat. Softw. 67, 1–48. doi: 10.18637/jss.v067.i01

732 Benito-Garzón, M., Alía, R., Robson, T. M., Zavala, M.A., 2011. Intra-specific variability

733 and plasticity influence potential tree species distributions under climate change.

734 Global Ecol. Biogeogr. 20, 766–778. https://doi.org/10.1111/j.1466-

735 8238.2010.00646.x

736 Benito-Garzón, M., Robson, T. M., Hampe, A., 2019. ΔTraitSDMs: species distribution

737 models that account for local adaptation and phenotypic plasticity. New Phytol. 222,

738 1757–1765. https://doi.org/10.1111/nph.15716

739 Bush, A., Mokany, K., Catullo, R., Hoffmann, A., Kellermann, V., Sgrò, C., McEvey, S.,

740 Ferrier, S., 2016. Incorporating evolutionary adaptation in species distribution

741 modelling reduces projected vulnerability to climate change. Ecol. Lett.19, 1468–

742 1478. https://doi.org/10.1111/ele.12696

743 Burgarella, C., Navascués, M., Zabal-Aguirre, M., Berganzo, E., Riba, M., Mayol, M.,

744 Vendramin, G.G., González-Martínez, S.C., 2012. Recent population decline and

745 selection shape diversity of taxol-related genes. Mol. Ecol. 21, 3006–3021.

746 https://doi.org/10.1111/j.1365-294X.2012.05532.x

747 Cedro, A., Cedro, B., 2015. Growth-climate relationships at yew and wild service trees

748 on th eastern edge of their range in Europe. Forest Syst., 24, e044.

749 http://dx.doi.org/10.5424/fs/2015243-07480.

750 De Araújo, C.B., Marcondes-Machado, L.O., Costa, G.C., 2014. The importance of biotic

751 interactions in species distribution models: A test of the Eltonian noise hypothesis

752 using parrots. J. Biogeogr. 41, 513–523. https://doi.org/10.1111/jbi.12234

753 Dolos, K., Bauer, A., Albrecht, S., 2015. Site suitability for tree species: Is there a positive

754 relation between a tree species' occurrence and its growth? Eur. J. For. Res. 134, 609–

755 621. https://doi.org/10.1007/s10342-015-0876-0

756 Dormann, C.F., Bobrowski, M., Dehling, D.M., Harris, D.J., Hartig, F., Lischke, H.,

757 Moretti, M.D., Pagel, J., Pinkert, S., Schleuning, M., Schmidt, S.I., Sheppard, C.S.,

758 Steinbauer, M.J., Zeuss, D., Kraan, C., 2018. Biotic interactions in species distribution

759 modelling: 10 questions to guide interpretation and avoid false conclusions. Global

760 Ecol. Biogeogr. 27, 1004–1016. https://doi.org/10.1111/geb.12759

761 Dubreuil, M., Riba, M., González-Martínez, S.C., Vendramin, G.G., Sebastiani, F.,

762 Mayol, M., 2010. Genetic effects of chronic habitat fragmentation revisited: Strong

763 genetic structure in a temperate tree, *Taxus baccata* (Taxaceae), with great dispersal

764 capability. Am. J. Bot. 97, 303–310. https://doi.org/10.3732/ajb.0900148

765 Dullinger, S., Gattringer, A., Thuiller, W., Moser, D., Zimmermann, N.E., Guisan, A.,

766 Willner, W., Plutzar, C., Leitner, M., Mang, T., Caccianiga, M., Dirnböck, T., Ertl, S.,

767 Fischer, A., Lenoir, J., Svenning, J.C., Psomas, A., Schmatz, D.R., Silc, U., Vittoz, P.,

768 Hülber, K., 2012. Extinction debt of high-mountain plants under twenty-first-century

769 climate change. Nat. Clim. Change 2, 619–622. https://doi.org/10.1038/nclimate1514

770 Ehrlén, J., Morris, W.F., 2015. Predicting changes in the distribution and abundance of

771 species under environmental change. Ecol. Lett. 18, 303–314.

772 https://doi.org/10.1111/ele.12410.

773 Elith, J., Franklin, J., 2013. Species distribution modeling, in: Levin, S.A., (Ed.),

774 Encyclopedia of Biodiversity (2nd Edition). Elsevier, pp. 692-705. ISBN

775 9780123847201. https://doi.org/10.1016/B978-0-12-384719-5.00318-X.

776 Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J.,

777 Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A.,

778 Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. McC. M., Peterson,

779 A.T., Phillips, S. J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberón, J.,

780 Williams, S., Wisz, M.S., Zimmermann, N.E., 2006. Novel methods improve

781   prediction of species' distributions from occurrence data. Ecography 29, 129–151.

782   https://doi.org/10.1111/j.2006.0906-7590.04596.x

783   Elith, J., Kearney, M., Phillips, S., 2010. The art of modelling range-shifting species.

784   Methods Ecol. Evol. 1, 330–342. https://doi.org/10.1111/j.2041-210X.2010.00036.x

785   Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E.,Yates, C.J., 2011. A statistical

786   explanation of MaxEnt for ecologists. Divers. Distrib. 17, 43–57.

787   https://doi.org/10.1111/j.1472-4642.2010.00725.x

788   Eriksson, Å., Eriksson, O., 2000. Population dynamics of the perennial *Plantago media*

789   in semi-natural grasslands. J. Veg. Sci. 11, 245–252. https://doi.org/10.2307/3236803

790   Fréjaville, T., Viscaíno-Palomar, N., Fady, B., Kremer., A., Benito-Garzón, M., 2020.

791   Range margin populations show high climate adaptation lags in European trees. Glob.

792   Change Biol. 26, 484-495. https://doi.org/10.1111/gcb.14881

793   García, D., Obeso, J.R., 2003. Facilitation by herbivore-mediated nurse plants in a

794   threatened tree, *Taxus baccata*: Local effects and landscape level consistency.

795   Ecography 26, 739–750. https://doi.org/10.1111/j.0906-7590.2003.03601.x

796   García, D., Zamora, R., Hódar, J.A., Gómez, J.M., Castro, J., 2000. Yew (*Taxus baccata*

797   L.) regeneration is facilitated by fleshy-fruited shrubs in Mediterranean environments.

798   Biol. Conserv. 95, 31–38. http://dx.doi.org/10.1016/S0006-3207(00)00016-1

799   Giannini, T.C., Chapman, D.S., Saraiva, A.M., Alves-dos-Santos, I., Biesmeijer, J.C.,

800   2013. Improving species distribution models using biotic interactions: a case study of

801   parasites, pollinators and plants. Ecography 36, 649–656.

802   https://doi.org/10.1111/j.1600-0587.2012.07191.x

803   Godsoe, W., Franklin, J., Blanchet, F.G., 2017. Effects of biotic interactions on modeled

804   species' distribution can be masked by environmental gradients. Ecol. Evol. 7, 654–

805   664. https://doi.org/10.1002/ece3.2657

806   González-Martínez, S.C., Dubreuil, M., Riba, M., Vendramin, G.G., Sebastiani, F.,

807   Mayol, M., 2010. Spatial genetic structure of *Taxus baccata* L. in the western

808   Mediterranean Basin: Past and present limits to gene movement over a broad

809   geographic scale. Mol. Phylogenet. Evol. 55, 805–815.

810   https://doi.org/10.1016/j.ympev.2010.03.001

811   Hijmans, R.J., Phillips, S., Leathwick, J., , Elith, J., 2017. dismo: Species Distribution

812   Modeling. R package version 1.1-4. https://CRAN.R-project.org/package=dismo

813   Holt, R.D., 2009. Bringing the Hutchinsonian niche into the 21st century: Ecological and

814   evolutionary perspectives. P. Natl. Acad. Sci. USA 106, 19659–19665.

815   https://doi.org/10.1073/pnas.0905137106

816   Hutchinson, G.E., 1957. Concluding remarks. Cold Spring Harb. Sym. 22, 415-427.

817   Iszkuło, G., 2010. Success and failure of endangered tree species: Low temperatures and

818   low light availability affect survival and growth of european yew (*Taxus baccata* l.)

819   seedlings. Pol. J. Ecol. 58, 259–271.

820   Iszkuło, G., Didukh, Y., Giertych, M.J., Jasińska, A.K., Sobierajska, K., Szmyt, J., 2012.

821   Weak competitive ability may explain decline of *Taxus baccata*. Ann. For. Sci. 69,

822   705–712.https://doi.org/10.1007/s13595-012-0193-4

823   Kenward, M., Roger, J., 1997. Small Sample Inference for Fixed Effects from Restricted

824   Maximum Likelihood. Biometrics 53, 983–997. doi:10.2307/2533558

825   Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest Package: Tests in

826   Linear Mixed Effects Models. J. Stat. Softw. 82, 1–26. doi: 10.18637/jss.v082.i13.

827   Larsson L., 2013. CooRecorder and Cdendro programs of the

828   CooRecorder/Cdendropackage version 7.6. http://www.cybis.se/forfun/dendro/

829 Lavabre, J.E., García, D., 2015. Geographic consistency in the seed dispersal patterns of

830 *Taxus baccata* L. in the Iberian Peninsula. Forest Syst., 24, e040.

831 http://dx.doi.org/10.5424/fs/2015243-07462.

832 Li, J., Lu, X., Cheng, K., Liu, W., 2020. StepReg: Stepwise Regression Analysis. R

833 package version 1.4.1. https://CRAN.R-project.org/package=StepReg

834 Lloret, F., Martinez-Vilalta, J., Serra-Diaz, J.M., Ninyerola, M., 2013. Relationship

835 between projected changes in future climatic suitability and demographic and

836 functional traits of forest tree species in Spain. Climatic Change 120, 449–462.

837 https://doi.org/10.1007/s10584-013-0820-6

838 Marcer, A., Méndez-Vigo, B., Alonso-Blanco, C., Picó, F.X., 2016. Tackling

839 intraspecific genetic structure in distribution models better reflects species

840 geographical range. Ecol. Evol. 6, 2084–2097. https://doi.org/10.1002/ece3.2010

841 Maroso, F., Vera, M., Ferreiro, J., Mayol, M., Riba, M., Ramil-Rego, P., Martínez, P.,

842 Bouza, C., 2021. Genetic diversity and structure of *Taxus baccata* from the

843 Cantabrian-Atlantic area in northern Spain: A guide for conservation and management

844 actions. Forest Ecol. Manag. 482, 118844.

845 https://doi.org/10.1016/j.foreco.2020.118844

846 Mayol, M., Riba, M., Cavers, S., Grivet, D., Vincenot, L., Cattonaro, F., Vendramin,

847 G.G., González-Martínez, S.C., 2020. A multiscale approach to detect selection in

848 nonmodel tree species: Widespread adaptation despite population decline in *Taxus*

849 *baccata* L. Evol. Appl. 13, 143–160. https://doi.org/10.1111/eva.12838

850 Meynard, C.N., Gay, P.E., Lecoq, M., Foucart, A., Piou, C., Chapuis, M.P., 2017.

851 Climate-driven geographic distribution of the desert locust during recession periods:

852 Subspecies' niche differentiation and relative risks under scenarios of climate change.

853 Glob. Change Biol. 23, 4739–4749. https://doi.org/10.1111/gcb.13739

854   Ninyerola, M., Pons, X., Roure, J.M., 2005. Atlas Climático Digital de la Península

855   Ibérica. Metodología y aplicaciones en bioclimatología y geobotánica. ISBN 932860-

856   8-7. Universidad Autónoma de Barcelona, Bellaterra.

857   Oney, B., Reineking, B., O'Neill, G., Kreyling, J., 2013. Intraspecific variation buffers

858   projected climate change impacts on *Pinus contorta*. Ecol. Evol. 3, 437–449.

859   https://doi.org/10.1002/ece3.426

860   Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of

861   species   geographic   distributions.   Ecol.   Model.   190,   231–259.

862   https://doi.org/10.1016/j.ecolmodel.2005.03.026

863   Pearman, P.B., D'Amen, M., Graham, C.H., Thuiller, W., Zimmermann, N.E., 2010.

864   Within-taxon niche structure: Niche conservatism, divergence and predicted effects of

865   climate   change.   Ecography   33,   990–1003.   https://doi.org/10.1111/j.1600-

866   0587.2010.06443.x

867   Pearson, R.G., Dawson, T.P., 2003. Predicting the impacts of climate change on the

868   distribution of species: Are bioclimate envelope models useful? Global Ecol.

869   Biogeogr. 12, 361–371. https://doi.org/10.1046/j.1466-822X.2003.00042.x

870   Peterson, M.L., Doak, D.F., Morris, W.F., 2019. Incorporating local adaptation into

871   forecasts of species' distribution and abundance under climate change. Glob.Change

872   Biol. 25, 775–793. https://doi.org/10.1111/gcb.14562

873   Piovesan, G., Saba, E.P., Biondi, F., Alessandrini, A., Di Filippo, A., Schirone, B., 2009.

874   Population ecology of yew (*Taxus baccata* L.) in the central Apennines: Spatial

875   patterns and their relevance for conservation strategies. Plant Ecol. 205, 23–46.

876   https://doi.org/10.1007/sl1258-009-9596-1

877   R Core Team, 2019. R: A language and environment for statistical computing. R

878   Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-

879   project.org

880   Razgour, O., Forester, B., Taggart, J.B., Bekaert, M., Juste, J., Ibáñez, C., Puechmaille,

881   S.J., Novella-Fernandez, R., Alberdi, A., Manel, S., 2019. Considering adaptive

882   genetic variation in climate change vulnerability assessment reduces species range loss

883   projections. P. Natl. Acad. Sci. USA 116, 10418–10423.

884   https://doi.org/10.1073/pnas.1820663116

885   Ruprecht, H., Dhar, A., Aigner, B., Oitzinger, G., Klumpp, R., Vacik, H., 2010. Structural

886   diversity of English yew (*Taxus baccata* L.) populations. Eur. J. For. Res. 129, 189–

887   198. https://doi.org/10.1007/s10342-009-0312-4

888   Sanz, R., Pulido, F., Nogués-Bravo, D., 2009. Predicting mechanisms across scales:

889   Amplified effects of abiotic constraints on the recruitment of yew *Taxus baccata*.

890   Ecography 32, 993–1000. https://doi.org/10.1111/j.1600-0587.2009.05627.x

891   Schurr, F.M., Pagel, J., Cabral, J.S., Groeneveld, J., Bykova, O., O'Hara, R.B., Hartig,

892   F., Kissling, W.D., Linder, H.P., Midgley, G.F., Schröder, B., Singer, A.,

893   Zimmermann, N.E., 2012. How to understand species' niches and range dynamics: A

894   demographic research agenda for biogeography. J. Biogeogr. 39, 2146–2162.

895   Serra, Ll., García-Martí, X., 2012. Distribución del Tejo (*Taxus baccata*) en España. In:

896   II International Workshop about *Taxus baccata* L, pp.11-35. Fundació d'estudis

897   superiors d'Olot.

898   Sinclair, S. J., White, M.D., Newell, G.R., 2010. How useful are species distribution

899   models for managing biodiversity under future climate changes? Ecol. Soc. 15: 8.

900   http://www.ecologyandsociety.org/vol15/iss1/art8/.

901 Singer, A., Johst, K., Banitz, T., Fowler, M.S., Groeneveld, J., Gutiérrez, A.G., Hartig,

902   F., Krug, R.M., Liess, M., Matlack, G., Meyer, K.M., Pe'er, G., Radchuk, V.,

903   Voinopol-Sassu, A.J., Travis, J.M.J., 2016. Community dynamics under

904   environmental change: How can next generation mechanistic models improve

905   projections of species distributions? Ecol. Model. 326, 63–74.

906   https://doi.org/10.1016/j.ecolmodel.2015.11.007

907 Svenning, J.C., Magård, E., 1999. Population ecology and conservation status of the last

908   natural population of English yew *Taxus baccata* in Denmark. Biol. Conserv. 88, 173–

909   182. https://doi.org/10.1016/S0006-3207(98)00106-2

910 Svenning, J.C., Skov, F., 2004. Limited filling of the potential range in European tree

911   species. Ecol. Lett. 7, 565–573. https://doi.org/10.1111/j.1461-0248.2004.00614.x

912 Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. Science 240, 1285–

913   1293. https://doi.org/10.1126/science.3287615

914 Swab, R.M., Regan, H.M., Matthies, D., Becker, U., Bruun, H.H., 2015. The role of

915   demography, intra-species variation, and species distribution models in species'

916   projections under climate change. Ecography 38, 221–230.

917   https://doi.org/10.1111/ecog.00585

918 Talluto, M.V., Boulangeat, I., Ameztegui, A., Aubin, I., Berteaux, D., Butler, A., Doyon,

919   F., Drever, C.R., Fortin, M.J., Franceschini, T., Liénard, J., Mckenney, D., Solarik,

920   K.A., Strigul, N., Thuiller, W., Gravel, D., 2016. Cross-scale integration of knowledge

921   for predicting species ranges: A metamodelling framework. Global Ecol. Biogeogr.

922   25, 238–249. https://doi.org/10.1111/geb.12395

923 Thuiller, W., Albert, C.H., Dubuis, A., Randin, C., Guisan, A., 2009. Variation in habitat

924   suitability does not always relate to variation in species' plant functional traits. Biol.

925   Lett. 6, 120–123. https://doi.org/10.1098/rsbl.2009.0669

926 Thuiller, W., Münkemüller, T., Schiffers, K.H., Georges, D., Dullinger, S., Eckhart, V.

927 M., Edwards, T.C., Gravel, D., Kunstler, G., Merow, C., Moore, K., Piedallu, C.,

928 Vissault, S., Zimmermann, N.E., Zurell, D., Schurr, F.M., 2014. Does probability of

929 occurrence relate to population dynamics? Ecography 37, 1155–1166.

930 https://doi.org/10.1111/ecog.00836

931 Thomas, P.A., Polwart, A, 2003. *Taxus baccata* L. J. Ecol. 91, 489–524.

932 https://doi.org/10.1046/j.1365-2745.2003.00783.x

933 Valladares, F., Matesanz, S., Guilhaumon, F., Araújo, M.B., Balaguer, L., Benito-Garzón,

934 M., Cornwell, W., Gianoli, E., van Kleunen, M., Naya, D.E., Nicotra, A.B., Poorter,

935 H., Zavala, M.A., 2014. The effects of phenotypic plasticity and local adaptation on

936 forecasts of species range shifts under climate change. Ecol. Lett. 17, 1351–1364.

937 https://doi.org/10.1111/ele.12348.

938 Van Echelpoel, W., Boets, P., Landuyt, D., Gobeyn, S., Everaert, G., Bennetsen, E.,

939 Mouton, A., Goethals, P.L.M., 2015. Species distribution models for sustainable

940 ecosystem development, in: Young-Seuk, P., Lek, S., Baehr C., Jørgensen, S.E. (Eds.),

941 Developments in Environmental Modelling. Elsevier, pp. 115-134. ISBN

942 9780444635365. https://doi.org/10.1016/B978-0-444-63536-5.00008-9.

943 Wiens, J.A., Stralberg, D., Jongsomjit, D., Howell, C.A., Snyder, M.A., 2009. Niches,

944 models, and climate change: assessing the assumptions and uncertainties. P. Natl.

945 Acad. Sci. USA 106, 19729–19736. https://doi.org/10.1073/pnas.0901639106

946 Wisz, M.S., Pottier, J., Kissling, W.D., Pellissier, L., Lenoir, J., Damgaard, C.F.,

947 Dormann, C.F., Forchhammer, M.C., Grytnes, J.A., Guisan, A., Heikkinen, R.K.,

948 Høye, T.T., Kühn, I., Luoto, M., Maiorano, L., Nilsson, M.C., Normand, S., Öckinger,

949 E., Schmidt, N.M., Termansen, M., Timmermann, A., Wardle, D.A., Aastrup, P.,

950 Svenning, J.C., 2013. The role of biotic interactions in shaping distributions and

951   realised assemblages of species: Implications for species distribution modelling. Biol.

952   Rev. Cab. Philos. 88, 15–30. https://doi.org/10.1111/j.1469-185X.2012.00235.x

953   Zimmermann, N.E., Edwards, T.C., Graham, C.H., Pearman, P.B., Svenning, J.C., 2010.

954   New trends in species distribution modelling. Ecography 33, 985–989.

955   https://doi.org/10.1111/j.1600-0587.2010.06953.x

Figures

Figure 1. Occurrences of *Taxus baccata* in the Iberian Peninsula used for modelling the species' distribution. Different colours indicate distinct adaptive groups, and locations used to measure tree growth (BAI5) are shown as black triangles. The four figures show the whole set of occurrences (a), and three of the nine occurrence datasets filtered by their predicted growth, i.e. including the 50% (b), 30% (c) and 10% (d) of the top predicted growth (BAI5) values (see text for further details). The number of occurrences is indicated in each case for both the *Continental* (C) and *Mild* (M) adaptive groups and for the species as a whole (SP).


42
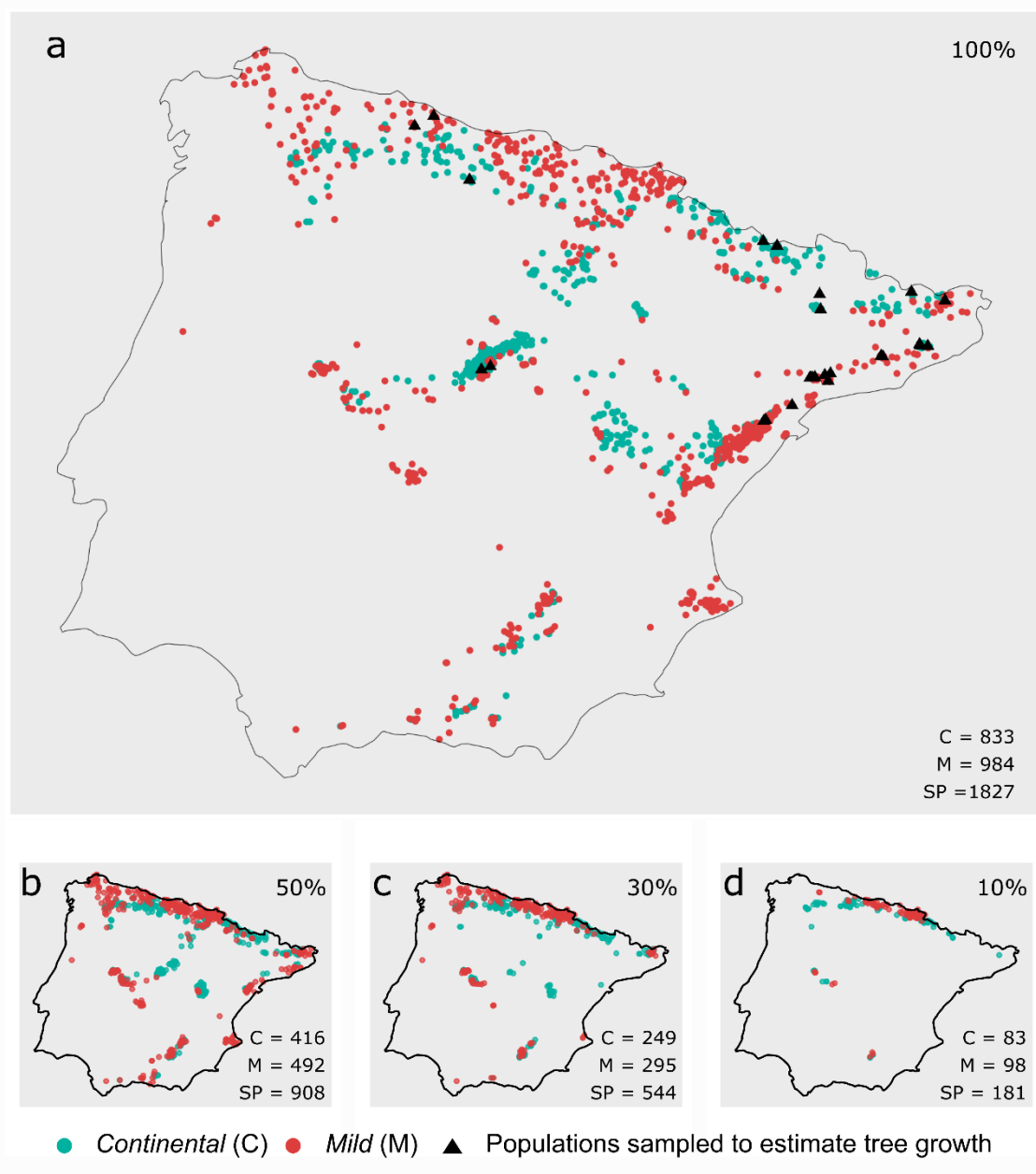
968   **Figure 2.** Scheme of the modelling approach implemented. Models were built following

969   three main criteria: 1) the type of predictors used: only climatic (*Clim*) and both climatic

970   and biotic predictors (*ClimBio*); 2) groups of occurrences: two adaptive groups

971   (*Continental* and *Mild*) and all occurrences without considering adaptive groups (SP); 3)

972   unfiltered (100% of the occurrences) or filtered datasets considering individual

973   performance according to predicted growth. In total, we generated 60 distribution models.

974   Twenty additional composite models were also obtained combining the results of

975   *Continental* and *Mild* distribution models (C+M). The predictive performance of all

976   models was assessed using AUC, as well as the correlation coefficient (Spearman's Rho)

977   between the predicted model suitability and projected growth. The best model for each

978   combination of predictor type and group of occurrences (i.e., same group and predictors

979   but different proportion of occurrences used) was selected by choosing those with the

980   highest suitability-growth correlation and with AUC $\geq$ 0.9. Range size was estimated both

981   for each of the best models and for those considering 100% occurrence data.
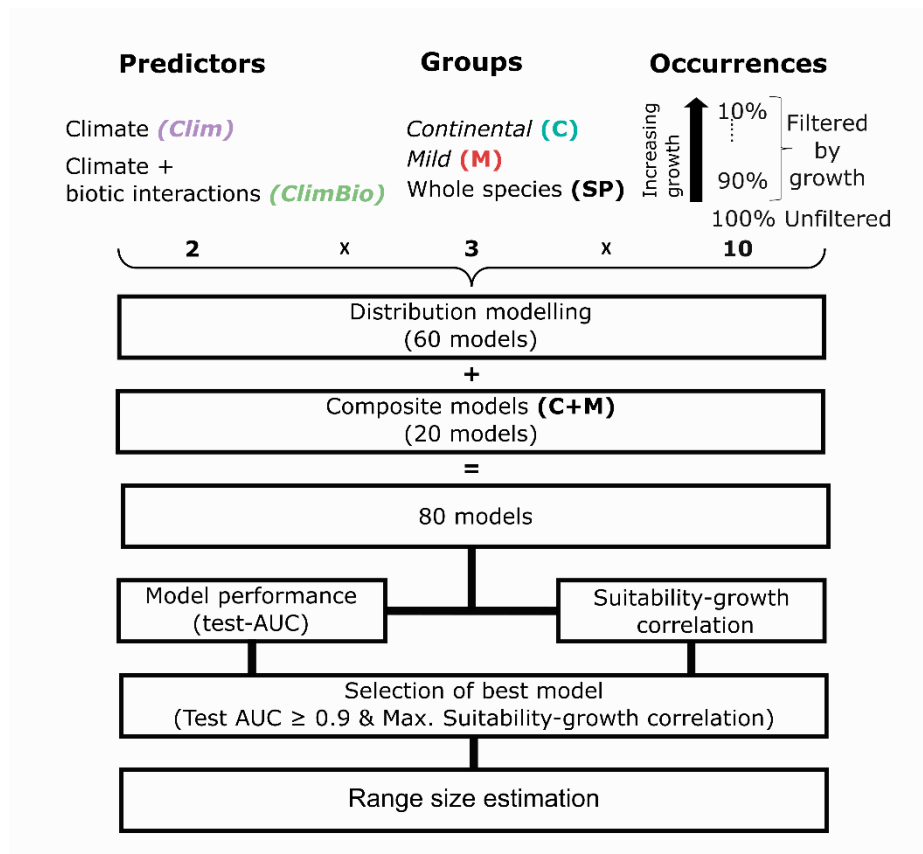
982

983

984   **Figure 3.** M Mean estimated tree growth values (BAI5: basal area increments during the
985   last five years) according to cover type. Bars are standard errors. UC: Uncovered; NP:
986   covered by needle-like perennials; BD: covered by broad-leaved deciduous; BP: covered
987   by broad-lived perennials; T: covered by conspecifics (*T. baccata*).
988
989

990

991
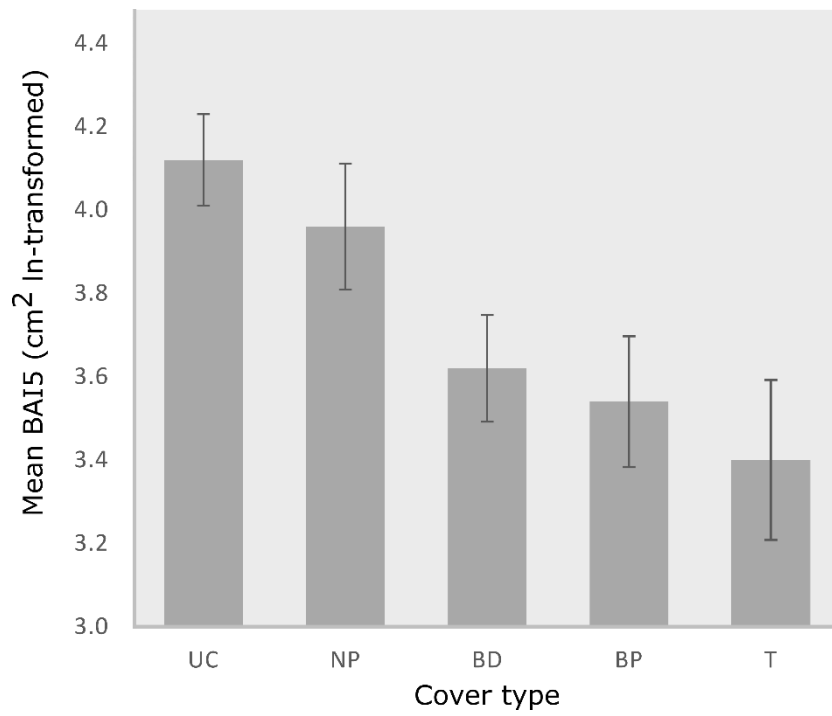
Figure 4. Projected suitability for unfiltered (i.e., using 100% of occurrences, top rows)
and the best models (i.e., using a proportion of the best growing occurrences which results
in a high AUC while optimizing the relationship between suitability and growth, bottom
rows), depending on the type of predictive variables (*Clim*, with only climatic predictors,
or *ClimBio*, with both climatic and biotic predictors) and the occurrence group
considered: *Continental* (C) or *Mild* (M) adaptive groups, the combination of both
adaptive groups (C+M), and without considering adaptive groups (SP). The percentages
in the plots indicate the percentage of top predicted growth values used to build the best
models (see details in Material & Methods).
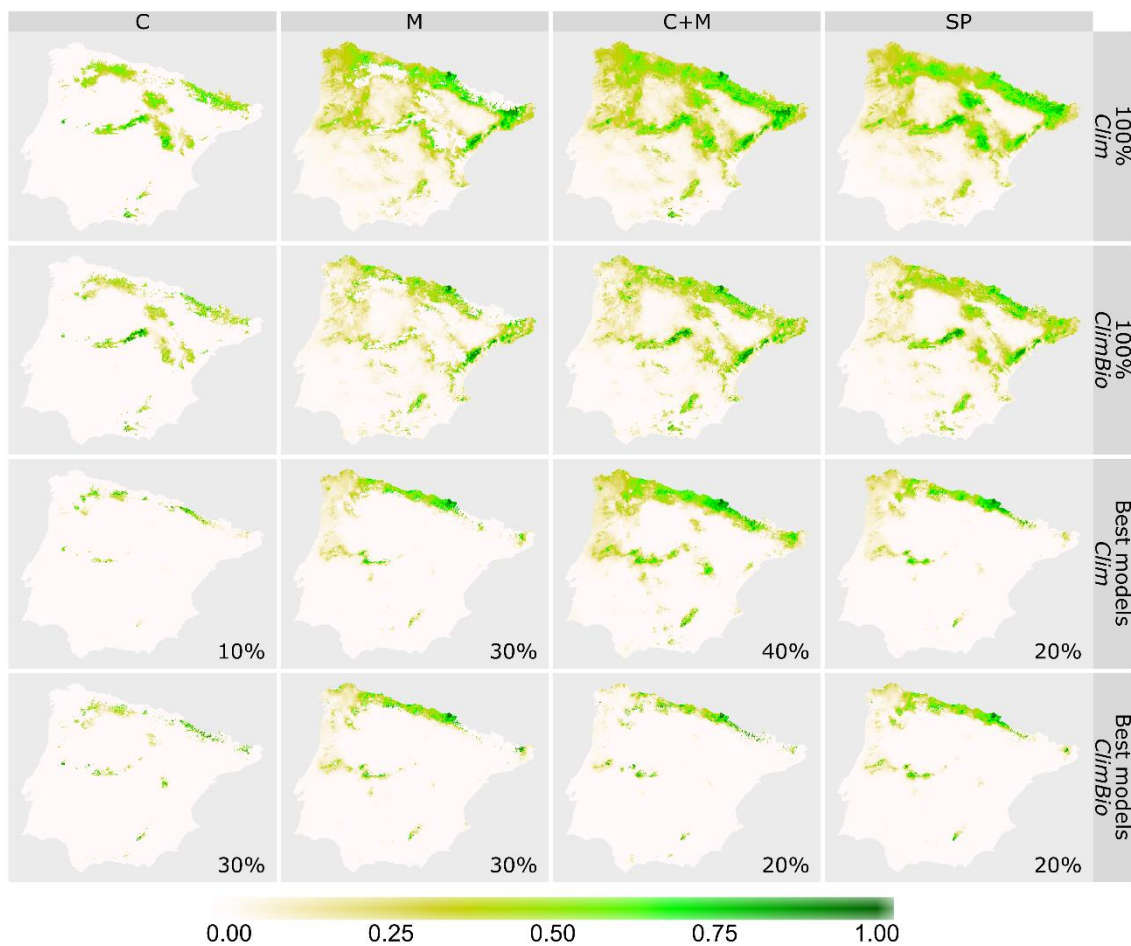
1004　**Figure 5.** Projected range size (km^2) for unfiltered (i.e., using 100% of occurrences, a)

1005　and the best models (i.e., using a proportion of the best growing occurrences which results

1006　in a high AUC while optimizing the relationship between suitability and growth, b),

1007　depending on the type of predictive variables (*Clim*, with only climatic predictors, or

1008　*ClimBio*, with both climatic and biotic predictors) and the occurrence group considered:

1009　*Continental* (C) or *Mild* (M) adaptive groups, the combination of both adaptive groups

1010　(C+M), and without considering adaptive groups (SP). The percentages in the plots

1011　indicate the percentage of top predicted growth values used to build the best models (see

1012　details in Material & Methods).

1013

1014
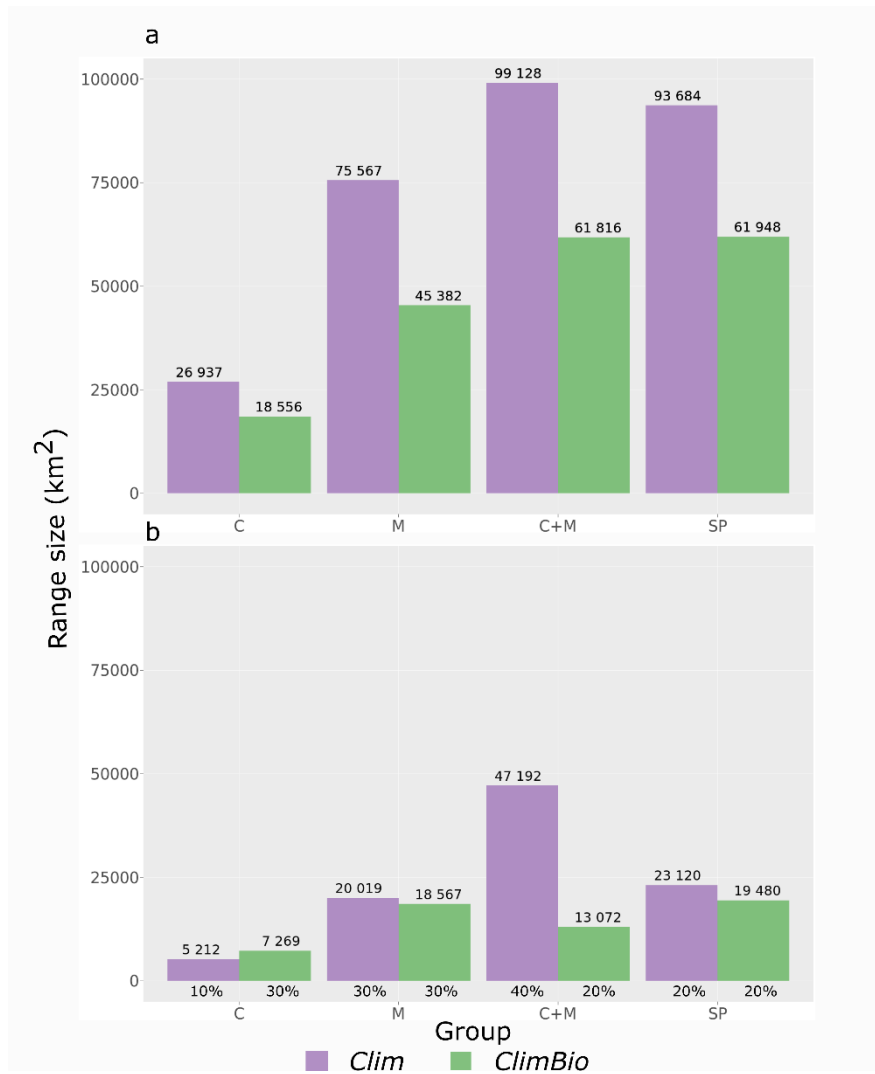
1015

1016

1017   **Figure 6.** Spearman's rank correlation coefficients between predicted suitability values

1018   obtained for each model and the predicted growth (BAI5) for each occurrence. The "X"

1019   axis indicates the percentage of occurrences used in each case (e.g., 100% refers to the

1020   unfiltered dataset with all the occurrence data, 90% is a filtered dataset including the 90%

1021   of the top predicted growth values, and so on). Models using only climate variables (*Clim*)

1022   are shown in lilac, while those based on both climate and biotic variables (*ClimBio*) are

1023   depicted in green. Significant ($p < 0.05$) and non-significant correlations are indicated by

1024   triangles and dots, respectively. C, models for the *Continental* adaptive group; M, models

1025   for the *Mild* adaptive group; C+M, combination of C and M model outputs; SP, models

1026   not considering the presence of adaptive groups.

1027
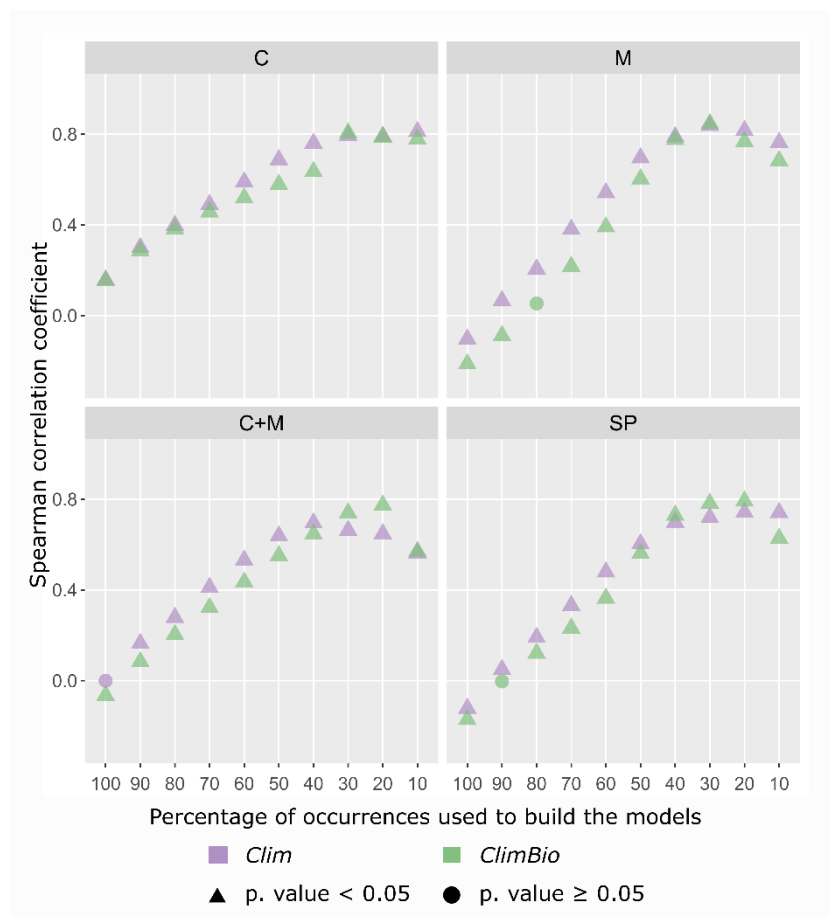
1028

1029 **Tables**

1030 **Table 1.** Location of the sampled populations used to assess individual performance
1031 (BAI5: basal area increments during the last five years, in cm^2). N: Number of sampled
1032 individuals; Census size: estimated number of reproducing individuals (Diameter at
1033 Breast Height > 2.5 cm); SpringP: spring precipitation (mm); AutumnT$_{max}$: autumn mean
1034 daily maximum temperature (^{o}C). Populations are ordered according to increasing
1035 SpringP. Coordinates are in decimal degrees in the World Geodetic System 1984
1036 (WGS84).

1037

Locality	Latitude	Longitude	N	Census size	SpringP	AutumnT$_{max}$	BAI5
Bruixa	41.294	0.869	4	150	159	17.2	24.6
Pas de l'Osca	42.039	0.840	14	125	168	15.1	25.8
Sant Jeroni	41.605	1.815	9	45	171	15.0	20.6
La Pena	41.355	1.086	12	300	180	16.9	22.7
Cardó	40.958	0.587	8	110	183	19.3	79.8
Agulles	41.610	1.788	8	75	185	16.3	19.0
Taverna	41.285	0.800	15	62	188	17.5	22.1
Vidalbar	41.288	0.831	26	250	193	17.1	30.5
Font Fresca	41.271	1.065	2	500	199	18.0	66.9
Rafalgarí	40.747	0.211	10	80	212	17.6	60.1
Titllar	41.330	1.006	16	600	215	17.3	60.8
Font del Teix	40.765	0.233	8	70	220	18.5	30.0
Espills	42.204	0.800	5	80	222	17.3	44.0
Canencia	40.873	-3.780	11	200	233	16.1	28.2
Font Negra	41.780	2.327	1	56	243	13.9	14.2
Orri	42.287	2.638	13	1000	270	16.8	35.1
Turó de l'Home	41.772	2.449	5	90	274	11.7	28.4
Rascafría	40.823	-3.900	6	400	292	14.7	57.9
Torrent de la Mina	41.766	2.345	13	160	300	12.6	38.4
Tosande	42.838	-4.552	10	2000	318	14.8	45.3
Estremera	42.343	2.136	4	4	320	13.4	32.8
Pineta	42.670	0.095	11	150	328	14.6	25.4
Bujaruelo	42.700	-0.117	7	4000	354	14.4	99.5
Penya Mayor	43.294	-5.504	8	4000	407	15.2	130.4
El Sueve	43.443	-5.253	9	8000	412	16.4	156.9

1038

1039

Table 2. Predictive performance for unfiltered models (i.e., including 100% of occurrences) and the best models, depending on the type of predictive variables (*Clim*, with only climatic predictors, or *ClimBio*, with both climatic and biotic predictors) and the occurrence group considered: *Continental* (C) or *Mild* (M) adaptive groups, the combined output of both adaptive groups (C+M), and without considering adaptive groups (SP). Occurrences – 100%, unfiltered models using the whole set of occurrences; 40%, 30%, 20%, 10%, filtered models including the 40%, 30%, 20% and 10% of the top predicted growth values. AUC – Area Under the Curve for test points. Correlation – Spearman's rank correlation coefficient between occurrence-derived suitability and predicted growth (BAI5) for each model.

Predictors	Group	Occurrences	AUC	Correlation	p-value
Clim	C	100%	0.970	0.157	<0.001
Clim	M	100%	0.886	-0.104	<0.01
Clim	C+M	100%	0.883	-0.001	0.968
Clim	SP	100%	0.880	-0.123	<0.001
ClimBio	C	100%	0.973	0.153	<0.001
ClimBio	M	100%	0.918	-0.213	<0.001
ClimBio	C+M	100%	0.916	-0.065	<0.01
ClimBio	SP	100%	0.912	-0.172	<0.001
Clim	C	10%	0.995	0.811	<0.001
Clim	M	30%	0.971	0.838	<0.001
Clim	C+M	40%	0.915	0.695	<0.001
Clim	SP	20%	0.973	0.742	<0.001
ClimBio	C	30%	0.983	0.805	<0.001
ClimBio	M	30%	0.958	0.846	<0.001
ClimBio	C+M	20%	0.970	0.774	<0.001
ClimBio	SP	20%	0.971	0.792	<0.001

1052
1053
1054
1055
1056
1057
1058
1059
1060

Table 3. Permutation importance of the predictive variables for unfiltered models (i.e., including 100% of the occurrences) and the best models, depending on the type of predictive variables (*Clim*, with only climatic predictors, or *ClimBio*, with both climatic and biotic predictors) and the occurrence group considered: *Continental* (C) or *Mild* (M) adaptive groups, and without considering adaptive groups (SP). Occurrences – 100%, unfiltered models using all the occurrences; 30%, 20%, 10%, filtered models including the 30%, 20% and 10% of the top predicted growth (BAI5) values. AP – Annual Precipitation. SuP – Summer Precipitation. PS – Precipitation Seasonality. MeWiT – Mean Winter Temperature. TS – Temperature Seasonality. The most important predictive variables, individually or jointly accounting for at least the 50% of the permutation importance, are shown in green and orange, indicating positive and negative effects on predicted suitability, respectively.

Predictors	Group	Occurrences	AP	SuP	PS	MeWiT	TS	*Quercus humilis*	*Quercus pyrenaica*	*Pinus pinaster*	*Pinus nigra*	*Fagus sylvatica*
Clim	C	100%	6.272	4.731	1.203	**85.884**	1.910					
Clim	M	100%	**38.445**	17.001	11.238	**27.866**	5.449					
Clim	SP	100%	**53.681**	13.373	13.610	8.065	11.271					
ClimBio	C	100%	1.335	3.511	1.453	**76.306**	2.492	2.722	1.216	1.932	1.225	7.808
ClimBio	M	100%	**17.919**	4.357	1.596	**18.032**	3.229	12.697	3.791	**24.695**	1.581	12.103
ClimBio	SP	100%	13.663	3.833	3.836	1.510	1.761	**14.954**	6.503	4.073	3.765	**46.103**
Clim	C	10%	34.364	0.362	0.165	**64.273**	0.835					
Clim	M	30%	**67.216**	2.399	13.268	14.495	2.623					
Clim	SP	20%	**76.887**	4.202	14.379	1.767	2.766					
ClimBio	C	30%	9.082	1.015	3.606	**71.149**	2.962	0.582	2.460	2.101	4.646	2.397
ClimBio	M	30%	**73.218**	0.546	1.194	18.693	1.527	1.808	0.151	0.979	0.499	1.386
ClimBio	SP	20%	**86.721**	0.974	3.428	0.310	2.844	1.991	0.350	1.209	1.604	0.569