

¹ Estimating process-based model parameters from species ² distribution data

³ **Victor Van der Meersch ***, **Isabelle Chuine** CEFÉ, Université de Montpellier, CNRS, EPHE, IRD, Montpelier, France

- ⁵ 1. Two main types of species distribution models are used to project species range shifts in future climatic conditions: correlative and process-based models. Although there is some continuity between these two types of models, they are fundamentally different in their hypotheses (statistical relationships *vs* cause-to-effect relationships) and their calibration methods (dependent *vs* independent of the species observed distributions).
- ¹⁰ 2. One of the main limitation to the use of process-based models is the difficulty to parameterize them for a very large number of species. Our aim was to calibrate process-based models in the same way as correlative models, i.e. using the geographic distributions of species. We investigated the feasibility of using an evolutionary algorithm (called covariance matrix adaptation evolution strategy, CMA-ES) to calibrate these models. This method is well established in some fields (robotics, aerospace research, ...), but has never been used, to our knowledge, in ecology, despite its ability to deal with very large space dimensions. Using tree species occurrence data across Europe, we adapted the CMA-ES algorithm to find appropriate values of model parameters. We estimated simultaneously 27 to 77 parameters of two process-based models simulating forest tree's ecophysiology for three species with varying range sizes and geographical distributions. We compared the performance of CMA-ES to a commonly used Approximate Bayesian Computation (ABC) method.
- ¹⁵ 3. CMA-ES provided parameter estimates leading to better prediction of species distribution than parameter estimates based on experts knowledge. It was more efficient than ABC, and provided better parameter sets for the same amount of computation time. Predictions of process-based models calibrated with CMA-ES were as good as predictions of more simple correlative models calibrated with standard optimisation algorithms. Our results also revealed that some model parameters and processes were strongly dependent, and different parameter combinations could therefore lead to high model accuracy.
- ²⁰ 4. CMA-ES is an efficient state-of-the-art method to calibrate process-based models with a large number of parameters using species occurrence data. Inverse modelling using CMA-ES is a powerful method to calibrate process-based parameters which can hardly be measured.

³¹ **Keywords:** calibration, evolutionary algorithm, cma-es, species distribution model, process-based model,
³² trees

³³ **Introduction**

³⁴ The speed and magnitude of projected climate changes are profoundly affecting species distributions, ecological communities and ecosystem processes, and numerous ecological systems are now approaching

*Corresponding author - victor.vandermeersch@cefe.cnrs.fr

36 tipping points (Lenton *et al.* 2008; Barnosky *et al.* 2012; Steffen *et al.* 2018). Large uncertainties on the
37 persistence and the resilience of ecosystems exist. Ecological forecasting has now become a critical tool
38 for managers and decision-makers (Urban 2015), and robust predictive approaches are necessary to pro-
39 vide reliable projections of species geographic range shifts and ecosystems functioning (Mouquet *et al.*
40 2015). Forecasting the dynamics of ecological systems for the upcoming decades and centuries is very
41 difficult, because ecological systems are extremely complex, influenced by a lot of factors and processes,
42 and climatic conditions with no analogous in the recent past are forecasted to become common (Williams
43 *et al.* 2007; Radeloff *et al.* 2015; Fitzpatrick *et al.* 2018). Ecological models have thus increased in com-
44 plexity over the last 50 years, incorporating more and more processes described with various degrees of
45 complexity depending on their objectives.

46 Nowadays, two main types of species distribution models (SDM) are used to project species range
47 shifts in future climatic conditions: correlative and process-based models (Dormann *et al.* 2012). The
48 vast majority of currently used SDMs are correlative: they seek to find statistical relationships between
49 various environmental descriptors and species presence and absence. They assume there is an equilib-
50 rium between species distribution and environment (equilibrium postulate, Guisan & Thuiller 2005), and
51 that species niche is stable over time (niche conservatism, Pearman *et al.* 2008). Most of them include
52 a fairly large number of predictors (particularly in machine-learning approaches), and consider flexible
53 transformations (linear, quadratic...) and interactions between them (Merow *et al.* 2014). Even though
54 some authors advocate for “putting more biology into SDMs” (Higgins *et al.* 2012), parameters have no a
55 priori defined ecological meaning (Dormann *et al.* 2012) and shape of response curves to environmental
56 variables is generally not constrained based on biological considerations. Although these models are not
57 always used correctly (Araújo *et al.* 2019; Santini *et al.* 2021), their flexibility makes them an important
58 tool in predictive ecology (Mouquet *et al.* 2015). They have been widely used especially to generate species
59 range projections under current and future climates (e.g. Guisan & Thuiller 2005). Nevertheless, their abil-
60 ity to accurately describe the effects of climate on species distributions has recently been questioned (e.g.
61 Fourcade *et al.* 2018; Journé *et al.* 2020; Warren *et al.* 2021). For all these reasons, another kind of models
62 has been developed. Process-based models aim to translate into mathematical equations our knowledge
63 about the physiological and ecological processes involved in an organism’s life, such as growth, reproduc-
64 tion, survival, movement, and interactions with other livebeings. Process-based models take more time to
65 develop and are more challenging to use, but they might provide greater comprehension of the complexity

66 of ecosystem dynamics and more robust projections in novel conditions (Evans 2012; Zurell *et al.* 2016;
67 Singer *et al.* 2016; Urban *et al.* 2016). A wide variety of process-based models exists, from quite simple
68 models (e.g. Kleidon & Mooney 2000) to much more complex ones (e.g. Dufrêne *et al.* 2005). They all rely
69 on an explicit representation of causal relationships, with a direct biological interpretation (Connolly *et*
70 *al.* 2017). The choices about the specific processes to include into the model are made based on theory,
71 empirical observations and the objectives of the research, and modeler subjectivity may play an important
72 role. One of the challenges is to build a model with the appropriate amount of complexity: a too simplistic
73 model might be unrealistic whereas a very complex model could be far beyond our ability to understand it
74 (because of interconnected mechanisms) and calibrate it. Each model relies on different hypotheses with
75 its own balance of complexity, accuracy and parsimony - and thus different numbers of unknown param-
76 eters to calibrate. Generally, the parameter values are obtained both from field or laboratory observations
77 and experiments made by the modelers themselves or already available in the literature, as well as from
78 statistical inference for each modeled process using process-specific data.

79 Calibration (i.e. parameter estimation) is a fundamental step in the modelling process. It fixes the model
80 in reality, and allow it to reproduce this reality with more or less success. The result of the calibration pro-
81 vides insights on the ability of the model to reproduce and explain the reality (model predictive power).
82 Calibration of complex models such as process-based SDMs is time-consuming, and modelers are often
83 challenged by the dimension of the parameter space, the complexity of the possible correlations among
84 parameters, and the scarcity of observed experimental data to calibrate them. Parameter inference can be
85 achieved through many methods which have been developed in the last decades. Most of them fall into
86 two categories: Bayesian inference or maximum likelihood estimation. On one hand, Bayesian inference
87 aims at estimating parameter posterior distribution while taking into account prior belief. On the other
88 hand, maximum likelihood methods aim at finding the parameters that maximize the model goodness of
89 fit, and are either deterministic (e.g. Nelder-Mead method) or stochastic (e.g. simulated annealing, evo-
90 lutionary algorithms). One of the promising method to solve high-dimensional optimization problems is
91 the Approximate Bayesian Computation (ABC), which is increasingly being used in ecological modelling
92 (Csilléry *et al.* 2010 ; Beaumont 2010). ABC combines a computational efficient approximation of pos-
93 terior distribution with the advantages of Bayesian parameter inference (including the quantification of
94 parameter uncertainty), and is thus particularly useful when numerical evaluation of the likelihood func-
95 tion is computationally not possible. Recently, an other approach belonging to the evolutionary algorithm

96 family, called Covariance Matrix Adaptation Evolution Strategy (CMA-ES), has been proposed (Hansen &
97 Ostermeier 2001). One of the advantage of CMA-ES is its ability to cumulate information over iterations
98 in order to adapt its own parameters (in particular the covariance matrix), which makes it more robust
99 to noise. CMA-ES is especially performant for non-separable problems (i.e. when the model parameters
100 are dependent) and large search space. This method has been successfully applied in various fields such
101 as aerospace (e.g. Collange *et al.* 2010), optics (e.g. Gagné *et al.* 2008), and robotics (e.g. Hill *et al.* 2020).
102 CMA-ES is acknowledged to be one of the most efficient approaches in continuous black-box optimization
103 (Hansen *et al.* 2010) but to our knowledge has never been used in ecology.

104 Here we explored the feasibility and interests of calibrating process-based SDMs with CMA-ES us-
105 ing species occurrence data as correlative SDMs do. We focused on two forest process-based models of
106 varying levels of complexity to evaluate the ability of CMA-ES to calibrate such models and compare its
107 performance relatively to the more widely used ABC method. The two models are PHENOFIT (27 to 36
108 parameters, Chuine & Beaubien 2001) and CASTANEA (77 parameters, Dufrêne *et al.* 2005). Each model
109 also emphasizes different ecological processes: while PHENOFIT focuses on phenology and how it relates
110 to survival and reproduction, CASTANEA focuses on carbon and water cycles. We aimed to assess the
111 ability of CMA-ES to calibrate these process-based models, to identify the limits of this method, and to
112 compare its efficiency with an ABC algorithm commonly used among ecologists. To this end, we cali-
113 brated the models in the same way as correlative models do (fitted process-based models *sensu* Dormann
114 *et al.* (2012)), that is to say, using the geographic distributions of species, and compare their efficiency to
115 that obtained with classical calibration, i.e. direct measurement of the parameters or statistical inference
116 of the different submodels using data for each modelled process. We focused on three European common
117 tree species, with different range extent and ecological preferences in order to evaluate the algorithms
118 performance in various geographical and climatic conditions. European beech (*Fagus sylvatica L.*) is one
119 of the most widely distributed broadleaved tree in Europe (from southern Sweden to Sicily and from Spain
120 to northwest Turkey), holm oak (*Quercus ilex L.*) is an evergreen broadleaved tree native of the Mediter-
121 ranean region, and silver fir (*Abies alba Mill.*) is a coniferous tree which mainly occurs in mountain forests
122 of Central Europe and some parts of Southern and Eastern Europe.

123 **1. Material and methods**

124 *1.1. Process-based models*

125 All versions of the models used for this study are coded in Java and distributed by the [CAPSIS](#) platform.

126 PHENOFIT is a process-based species distribution model for forest tree species which focuses on phe-
127 nology. It relies on the principle that the distribution of a tree species depends mainly on the synchro-
128 nization of its timing of development to the local climatic conditions ([Chuine & Beaubien 2001](#)). It is
129 composed of several submodels, including phenology models for leaves, flowers and fruits, and stress re-
130 sistance models. It simulates the fitness (survival and reproductive success) of an average individual using
131 daily meteorological data, soil water holding capacity and species specific parameters (see [Appendix A](#) for
132 details). PHENOFIT has been validated for several North American and European species by comparing
133 their known distribution to the modelled fitness (e.g. [Morin et al. 2007](#); [Saltré et al. 2013](#); [Duputié et al.](#)
134 [2015](#); [Gauzere et al. 2020](#)).

135 CASTANEA is an ecophysiological process-based model which simulates carbon and water fluxes in
136 forests ([Dufrêne et al. 2005](#)). The model simulates the ecosystem as an average tree with six compartments
137 (leaves, branches, stem, coarse roots, fine roots and reserves). It is much more complex than PHENOFIT,
138 with several processes described and computed, such as photosynthesis, stomatal opening, maintenance
139 and growth respiration, transpiration, and carbon allocation (see [Appendix A](#) for details). CASTANEA
140 requires daily meteorological variables and soil characteristics. The model has been initially validated at
141 stand scale for beech ([Davi et al. 2005](#)), and was then successfully applied to other European species (e.g.
142 [Davi et al. 2006](#); [Delpierre et al. 2012](#); [Davi & Cailleret 2017](#)).

143 *1.2. Data for the calibration*

144 *1.2.1. Climate and soil data*

145 Raw climatic variables were extracted from ERA5-Land hourly dataset ([Muñoz Sabater 2019, 2021](#)) from
146 1970 to 2000, at a spatial resolution of 0.1 degree in latitude and longitude. We calculated the daily mean
147 values of the following variables used by PHENOFIT and CASTANEA: minimum, mean and maximum
148 daily temperatures, mean dewpoint temperature, daily precipitation, daily global radiation and daily mean
149 wind speed. We computed the daily relative humidity with the ratio of vapor pressure and saturation vapor

150 pressure (both calculated with Clausius–Clapeyron equation) using *humidity* R package (Cai 2019). Daily
151 potential evapotranspiration was calculated with Penman–Monteith equation (FAO standard of hypothet-
152 ical grass reference surface) using a slightly modified version of the *ET()* function in *Evapotranspiration* R
153 package (Guo *et al.* 2016).

154 Water content at field capacity and wilting point data were extracted from EU-SoilHydroGrids (Tóth
155 *et al.* 2017) which is at 1km resolution. Percentage of sand, silt and clay particles, percentage of coarse
156 fragments, bulk density and soil depth were extracted from SoilGrids250m (Hengl *et al.* 2017) at a 250m
157 resolution. These data (except for soil depth) are provided at seven soil depths, so we summarized them
158 (weighted sum or weighted mean) taking into account each layer width and total soil depth. Finally, all
159 variables were upscaled at the ERA5-Land spatial resolution 0.1°.

160 *1.2.2. Tree occurrences in Europe*

161 Sources of occurrence data are known to differ even for common European trees (Duputié *et al.* 2014) and
162 this makes it quite challenging to gather comprehensive data at a sufficient spatial resolution all over Eu-
163 rope. The occurrence data we used essentially rely on the EU-Forest dataset (Mauri *et al.* 2017) which ben-
164 efits from inventory and monitoring programmes implemented in most European countries. As EU-Forest
165 is limited to forest ecosystems, we completed it with presence records extracted from the Global Biodiver-
166 sity Information Facility (GBIF 2022, see Appendix B for all download links) but removing observations
167 outside natural species ranges as defined by Atlas Flora Europeae (AFE, Jalan & Suominen 1972–2005) and
168 EuroVegMap (Bohn *et al.* 2003). By doing so, we also included occurrences of isolated native trees living
169 outside forests, excluding records from arboreta or gardens where the species would have been planted as
170 an exotic. For holm oak, we also added occurrence records in the Mediterranean Basin from the WOODIV
171 database (Monnet *et al.* 2021), leaving out EU-Forest and GBIF records we had already gathered. We up-
172 scaled all species records at the ERA5-Land resolution (i.e. 0.1° cell, see 1.2.1. Climate data). We finally
173 obtained 21458 occurrence cells for beech, 6653 for holm oak and 5385 for silver fir (see Appendix B for
174 details).

175 All the datasets described above are presence-only data. Therefore, we generated cells where species
176 are supposed to be absent, i.e. pseudo-absence cells. In order to avoid as far as possible creating false
177 absence data, we used EU-Forest cells where the species is not reported present as pseudo-absence cells.
178 We assumed that national forest inventories were exhaustive (which is not true since only specific forest

179 plots in a 0.1° cell are monitored). We obtained 25423 absence cells for beech, 37931 for holm oak and 38365
180 for silver fir (see [Appendix C](#)).

181 We selected subsets of 2000 points (1000 presences and 1000 pseudo-absences) in order to reduce com-
182 putational costs. For each species, we generated ten presence clusters of similar bioclimatic conditions
183 based on annual climate normals computed with R package *dismo* ([Hijmans et al. 2021](#)) and ERA5-Land
184 variables. In each cluster, we randomly sampled a number of cells where the species is present propor-
185 tional to the total number of a number of cells where the species is present in the cluster. The aim of this
186 stratified random sampling was to make sure that all species environmental preferences were proportion-
187 ally represented. We then randomly sampled the same number of pseudo-absence cells (see [Appendix B](#)
188 for details).

189 *1.3. Model calibration*

190 *1.3.1. Covariance Matrix Adaptation Evolution Strategy principles*

191 Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is widely accepted as a robust optimization
192 algorithm for non-linear, non-convex, as well as non-separated optimization problems in continuous do-
193 main ([Hansen & Ostermeier 1996](#); [Hansen & Ostermeier 2001](#); [Hansen 2006](#)). It is based on the principle of
194 evolutionary biology, via recombination, mutation and selection of the most fit individuals (i.e. parameter
195 sets providing the best predictions). At each iteration:

196 - λ individuals are evaluated, i.e. model runs λ times with λ different parameter sets and the objective
197 function is evaluated

198 - the best μ individuals are selected

199 - the weighted mean individual m is computed (mean of the best μ parameter sets weighted by their ob-
200 jective function value) - covariance matrix C and step size σ are updated (with information accumulated
201 over several consecutive iterations)

202 - new λ individuals are sampled in a normal distribution $\mathcal{N}(m, \sigma C)$, with both recombination (via the
203 favorite solution m) and mutations (via the perturbations σC)

204 One of the strength of this approach lies in the combination of rank- μ -update, where prior information
205 from previous generations is exploited (mean of the previous covariance matrices, with a higher weight
206 for recent generations), and cumulation, where correlations between generations are retained in an evo-

207 lution path (sum of consecutive steps), to update the covariance matrix at each step (see Hansen 2016 for
208 a detailed description of the algorithm).

209 *1.3.2. CMA-ES in practice*

210 One of the advantages of CMA-ES is that it does not require a complex parameter tuning: as best parameter
211 values at a given time of the optimization process might no longer be efficient later, CMA-ES implements an
212 internal adaptation of its parameters. We only chose the population size λ , depending on the optimization
213 problem complexity (μ was set to $\lambda/2$). The default recommended value for λ is $4 + 3\ln(N)$, where N is
214 the number of parameters to calibrate (i.e. $\lambda \in [14, 17]$ in our case). We set $\lambda = 20$, in order to improve
215 the global search capability (Hansen & Kern 2004) and take advantage of the computation power at our
216 disposal. All model parameters were linear scaled into $[0; 10]$ so that the same standard deviation can
217 be applied to all parameters: here we chose $\sigma = 2$ (see [Nikolaus Hansen personal website](#) for practical
218 hints on variable encoding). Our stopping criterion for the optimization procedure was the budget, i.e. the
219 number of model runs.

220 For an easier use and the sake of reproducibility, we chose to use a pure R implementation of CMA-
221 ES available in the R package *cmaes* (Trautmann *et al.* 2011). The function *cma_es()* enables us to do λ
222 function evaluations in parallel so as to substantially reduce computation time. It also allows us to define
223 lower and upper bound constraints, by penalising individual fitness (i.e. objective function value) if it
224 violates the boundaries. We customized the *cma_es()* function to add an option to define death penalty
225 constraints (rejection of the infeasible individual who is sampled again), in order to define a range of
226 ecologically possible solutions in terms of inequality constraints between parameters (see [Appendix D](#) for
227 details about boundaries and constraints handling). Death penalty is the easiest way to handle constraints
228 when the feasible region is fairly large, but it is not perfect as there is no use of information from infeasible
229 points.

230 In the spirit of species distribution modelling, the objective function for the calibration was the area
231 under the receiver operating characteristic curve (AUC), evaluated against a subsets of 2000 points (see
232 [1.2.2. Tree occurrences](#)). Although AUC has been criticized as an imperfect measure of model performance
233 (Lobo *et al.* 2008; Leroy *et al.* 2018), we used it as objective function because our goal here was only to
234 calibrate models by maximizing discriminating capacity (i.e. potential to correctly classify presences and
235 absences) with a threshold-independent measure. We used the *AUC* R package (Ballings & Van den Poel

Table 1: Summary of model calibration settings. Average runtime was assessed on the GenOuest cluster.

Model	Output variable of interest	Number of parameters calibrated	Population size λ	Number of cores	Memory	Average runtime
PHENOFIT	Fitness index	[27; 36]	20	40	80 Go	~ 24 hours
CASTANEA	Carbon reserves	77	20	100	120 Go	~ 20 days

236 2013), and chose the two following model output variables as proxies of classification probabilities (i.e. used
 237 to determine if the species can be present or not): fitness index for PHENOFIT and carbon reserves for
 238 CASTANEA (see Appendix A).

239 We implemented the CMA-ES calibration on two computing clusters: GenOuest from IRISA-INRIA
 240 (genouest.org) and TGCC (*Très Grand Centre de Calcul*) from CEA (hpc.cea.fr). As the models are coded
 241 in Java (see 1.1 Process-based models), they need a process of deallocating memory handled by a *garbage*
 242 *collector*. For PHENOFIT, each function evaluation (i.e. each model simulation) was run on a 2-core com-
 243 puting unit in order to have enough computing resources for both simulation and garbage collection. We
 244 thus needed twice as many cores as functions evaluated in parallel. CASTANEA model requires a fairly
 245 high computation time, so we used a nested parallelism distribution, where each parallel simulation was
 246 distributed on 4 computing units. We thus used 4 times as many cores as functions evaluated in paral-
 247 lel, plus some extra cores for garbage collection. We used R package *future* (Bengtsson 2021) for parallel
 248 processing.

249 Regarding PHENOFIT model, we calibrated ten times each species parameter set using CMA-ES algo-
 250 rithm, with 5 repetitions on 2 random subsets of presences/pseudo-absences (see 1.2.2. Tree occurrences),
 251 except for beech. In the latter case, we ran 10 repetitions on 10 subsets (i.e. 100 calibrations) to investigate
 252 both the effect of subsampling and the effect of stochasticity on the calibration performance of CMA-ES.
 253 Since CASTANEA computing time was much higher (see Table 1), we ran only two calibration for each
 254 species (on 2 different random subsets).

255 1.3.3. Approximate Bayesian Computation

256 In order to provide elements of comparison between CMA-ES and an other widely used and effective
 257 optimization algorithm, we also implemented an Approximate Bayesian Computation (ABC) method. We

258 chose ABC because it is well established in ecology (e.g. Hartig *et al.* 2014; Lagarrigues *et al.* 2015; Vaart *et*
259 *al.* 2015; Gardner *et al.* 2020). Several versions of ABC exist, but here we only explore the most accessible
260 version, simple rejection. We chose this version because it does not require a tedious parameter tuning,
261 and it is therefore as easy as CMA-ES to apply to various calibration problems.

262 ABC rejection involves running the model a large number of times, with parameters randomly sampled
263 from their prior distributions, and then accepting the simulations closest to the observed data. These
264 retained simulations are then used to draw the posterior distributions of the model's parameters given the
265 data. We used uniform priors with the same lower and upper bounds than CMA-ES. We ran the model the
266 same number of times as CMA-ES (i.e. same number of objective function evaluations) and we obtain the
267 same computation time. Therefore, we had a fair comparison of the performance of the two algorithms.
268 We ran 30 ABC calibrations (10 repetitions on 3 subsets) of PHENOFIT for beech in order to be able to
269 make a robust comparison with CMA-ES. To avoid getting a beech specific comparison, we also ran 10
270 ABC calibrations (5 repetitions on 2 subsets) of PHENOFIT for holm oak (see [Appendix E](#)).

271 2. Results

272 2.1. Calibration results

273 Calibrations using species distribution data, either CMA-ES or ABC, are thereafter called backward cal-
274ibrations, and calibrations based on expert knowledge, observations and measurements of the processes
275 modelleds are called forward calibrations.

276 CMA-ES calibration of PHENOFIT model allows an average 17.2% increase of AUC across the three
277 species compared to forward calibration ([Figure 1](#)). The maximum increase is obtained for silver fir, from
278 0.72 to 0.9 (25%).

279 CMA-ES calibration of CASTANEA allows an average 23.7 % increase of AUC compared to forward cali-
280 bration ([Figure 2](#)), and a maximum increase obtained for holm oak (34.7%).

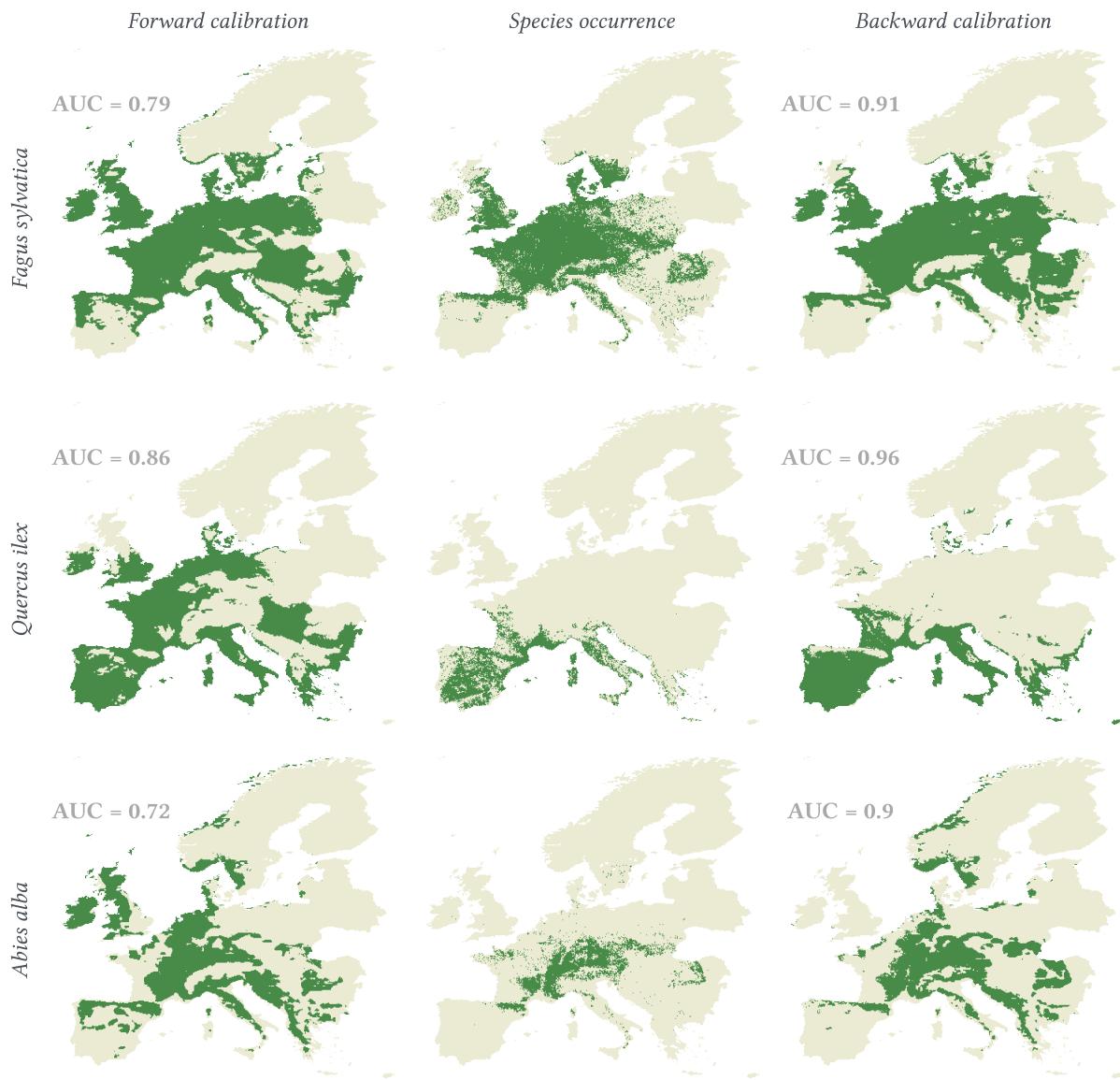


Figure 1: Species distribution maps obtained with PHENOFIT forward and backward calibrations, compared with observed species occurrences. Optimal threshold to dichotomize model predicted fitness index in presence/absence is the Youden index-based cut-off point. Note that models predict species climatic niche which is larger than the realized niche that corresponds to species presence map.

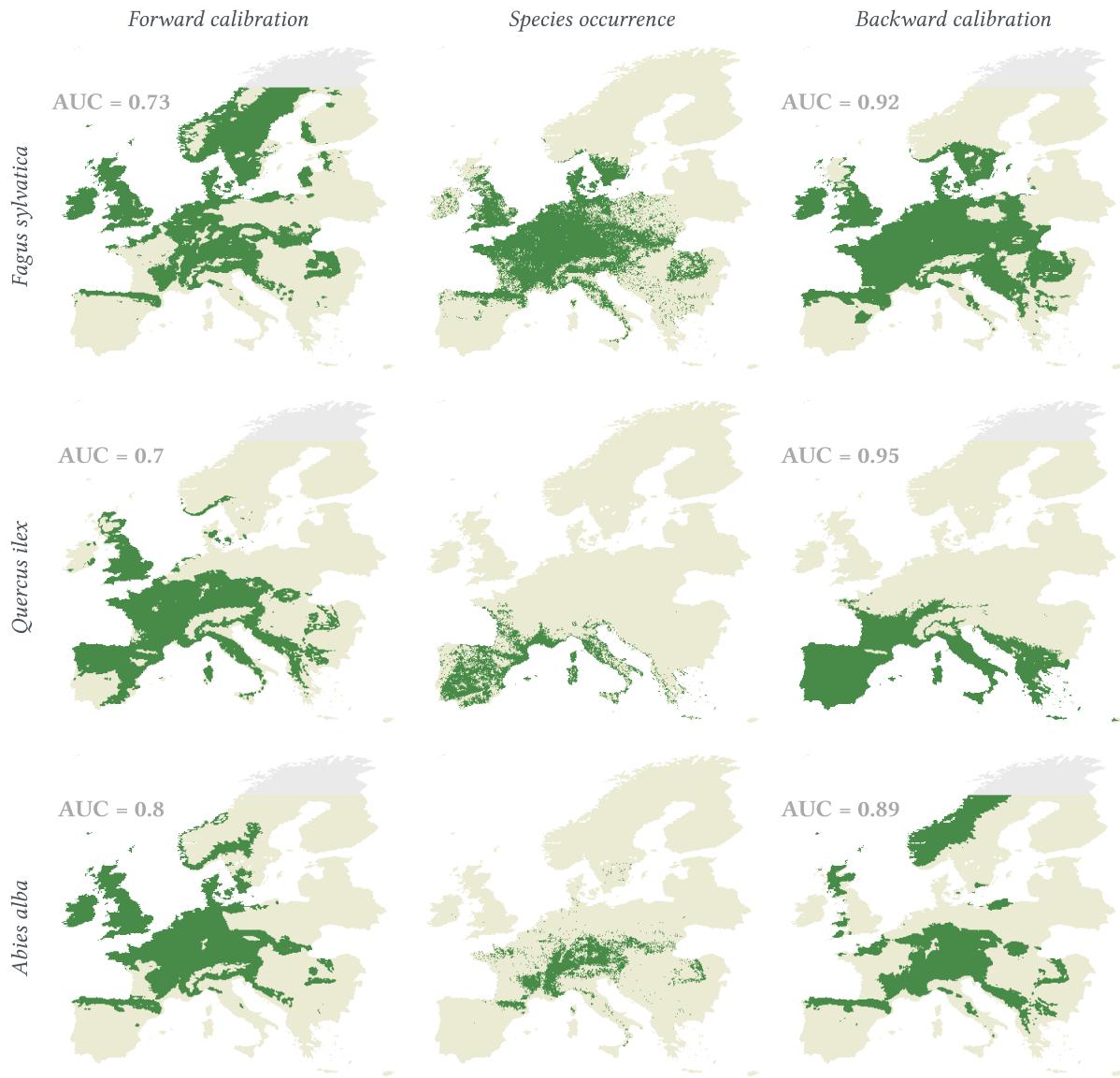


Figure 2: Species distribution maps obtained with CASTANEA forward and backward calibrations, compared with observed species occurrences. Optimal threshold to dichotomize model predicted carbon reserves in presence/absence is the Youden index-based cut-off point. Note that models predict species climatic niche which is larger than the realized niche that corresponds to species presence map.

281 2.2. Impacts of subsampling and calibration stochasticity

282 2.2.1. Variability of calibration performance

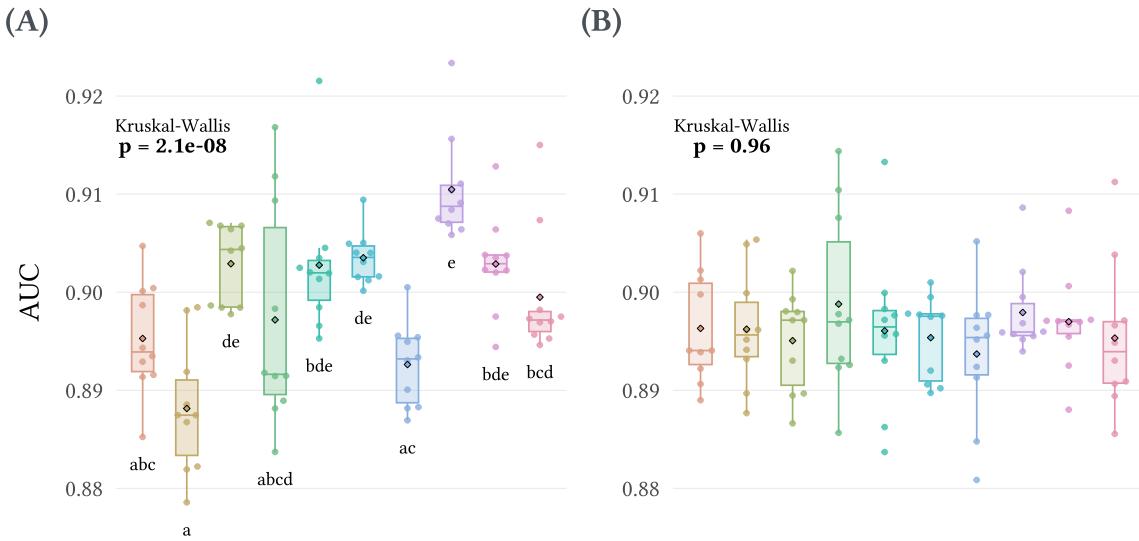


Figure 3: Effects of data sub-sampling and stochasticity on CMA-ES calibration using the PHENOFIT model for beech: **(A)** calibration AUC (calculated only with calibration cells) and **(B)** total AUC (calculated with all presence/absence cells). Each color is a different sub-sampling of occurrence data, each point is a calibration run. Diamonds (with black border) are mean AUC values. On **(A)**, the grouping letters represent the multiple comparisons with pairwise Dunn's tests.

283 The 100 calibrations of the PHENOFIT model realized for beech showed that random data subsampling
 284 had an effect on the final objective function value (i.e. the AUC computed on the 2000 calibration points).
 285 Kruskal-Wallis test was significant ($p = 2.1e-08$), meaning that at least one subset provided better AUC
 286 during calibration. According to Dunn's tests, 11 pairwise comparisons out of 45 were significant (Figure
 287 3.A.). The calibration AUC ranged from 0.879 to 0.923 over all subsets, with a mean value of 0.9.

288 However, more importantly, the repetition of calibrations on different subsets had no significant im-
 289 pact on the total AUC computed on all presence/absence points (Kruskal-Wallis test, $p = 0.96$). Thus, no
 290 subset led to an overall better prediction of the species distribution (see Figure 3.B.). The total AUC ranged
 291 from 0.881 to 0.914, with a mean value of 0.896.

292

293 2.2.2. Non-identifiability of parameters

294 We found a high variability in the parameter estimates of the leaf unfolding date submodel of PHENOFIT
 295 after 100 calibrations (Figure 4.A.). For example, the critical amount of chilling C_{crit} required to break bud
 296 dormancy and the critical amount of forcing F_{crit} required to break bud ranged from 1.02 to 149.96 and from
 297 1.5 to 79.26 respectively, with a mean value of 51.52 and 38.78. Their coefficient of variations were 126.7%
 298 and 51.9% respectively. Kendall correlation coefficient between C_{crit} and the threshold temperature of the
 299 response function to temperature during dormancy T_b is 0.64 ($p < 0.001$). Kendall correlation coefficient
 300 between F_{crit} and the mid-response temperature T_{50} is -0.55 ($p < 0.001$).

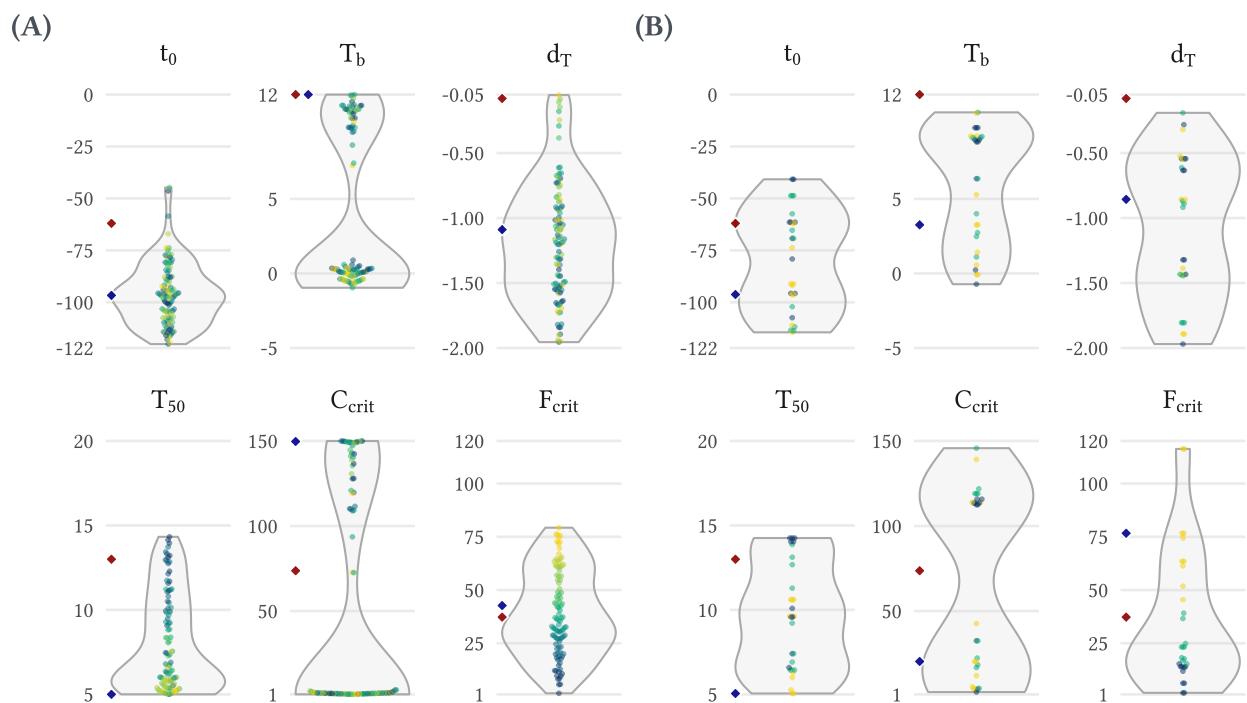


Figure 4: Effects of stochasticity of (A) CMA-ES and (B) ABC calibrations on PHENOFIT leaf unfolding model parameter values for beech. Y-axis limits are lower and upper bounds used during calibration. Each point is a calibrated parameter value, color gradient is based on F_{crit} values. Red diamonds are parameter values obtained with classical (forward) calibration, blue ones are parameter values obtained with the best backward calibration.

301 2.3. Comparing CMA-ES and ABC efficiency

302 The mean calibration AUC and total AUC obtained for beech with CMA-ES (0.899 and 0.896 respectively)
 303 were only slightly higher than those obtained with ABC (0.872 and 0.869). The best total AUC obtained
 304 with CMA-ES was 0.913, a higher value than the best solution obtained with ABC, 0.876. According to

305 Mann–Whitney tests, the distributions of AUC measures between the two methods with the three different
 306 occurrence subsets differed significantly (Figure 5). The outperformance of CMA-ES compared with ABC
 307 was the same for holm oak (see Figure E.1. in Appendix E). Both methods result in a high variability in the
 308 parameter estimates (Figure 4) .

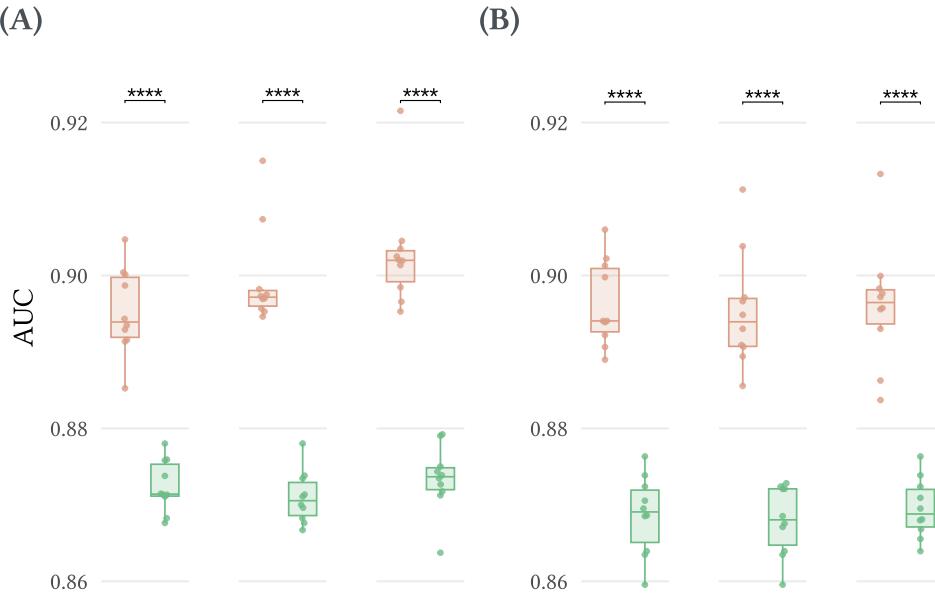


Figure 5: Comparison of CMA-ES and ABC-rejection methods, with three occurrence subsets of beech, on (A) calibration AUC (only calibration points) and (B) total AUC (every presence/absence points). Each point is a calibration run. The black horizontal bars represent the pairwise Mann–Whitney tests between the two methods on the same subset. CMA-ES is red and ABC green.

309 3. Discussion

310 3.1. Performance and advantages of CMA-ES to calibrate complex models in ecology

311 Our results are the first to demonstrate that inverse calibration with CMA-ES is feasible and provide good
 312 results for complex and runtime-expensive ecological models.
 313 With a subsampling of species occurrence data, the algorithm succeeds in finding parameter sets which
 314 provide high AUC values. The predictions of the calibrated models are sharply improved compared to
 315 classical (forward) parametrization (Figure 1 and Figure 2). AUC values obtained using CMA-ES were as
 316 good as the ones generally obtained with correlative models. Two striking examples are the increase in the
 317 performance of PHENOFIT model for silver fir, from 0.72 to 0.9, and of CASTANEA model for holm oak,
 318 from 0.7 to 0.95. Moreover, CMA-ES performed equally well regardless of the species occurrence subset

319 used during calibration (Figure 3.B.), and thus permitted to find a good compromise between computational
320 cost and calibration efficiency.

321 CMA-ES is a “generic” optimizer which can be applied to various problems. It is easy to use as it does
322 not require an extensive tuning to efficiently explore the parameter space. We only had to choose the
323 population size λ , and the initial search region (initial starting point and step size σ). As well as being
324 quasi parameter-free, CMA-ES has several structural advantages:

- 325 - the covariance matrix allows to learn second-order information (pairwise dependencies between param-
326 eters)
- 327 - the covariance matrix adaptation is particularly efficient to deal with ill-conditioned and non-separable
328 problems (Hansen *et al.* 2011)
- 329 - the update of the step size σ (i.e. mutation force) prevents premature convergence (Hansen & Ostermeier
330 2001)

331 CMA-ES has been shown to outperform several other optimization algorithms (Hansen *et al.* 2010), and is
332 usually the most efficient method when the target cost (i.e. the number of objective function evaluations)
333 is about $100 * N$ (N being the dimension of the parameter search space, Bäck *et al.* 2013). Here we show
334 that it outperforms the widely used ABC optimization method too.

335 *3.2. Non-identifiability of parameter values*

336 There can be a strong dependence between process-based model parameters.
337 In both models used in this study, biological mechanisms are explicitly calculated in several submodels
338 (e.g. a leaf unfolding submodel or a stomatal opening submodel). A submodel output has inevitably a sig-
339 nificant influence on the other submodels as biological processes can be highly dependent with feedbacks:
340 in CASTANEA, for example, the stomatal opening affects the photosynthesis, and *vice versa*. Within each
341 submodel, parameters are also strongly dependent because of structural correlations. To illustrate this
342 problem, we focused on beech leaf unfolding submodel of PHENOFIT (see Appendix A). This model has
343 6 parameters (Chuine 2000): a starting date of the processes (t_0), one parameter describing the response
344 function to temperature during the dormancy phase (T_b), two parameters describing the response function
345 to temperature during the phase of bud growth (d_T, T_{50}), and two parameters representing the sums of
346 the daily responses to temperature during bud dormancy (C_{crit}) and during bud growth (F_{crit}) that re-
347 spectively determine the date of bud dormancy break and the date of leaf unfolding (see Appendix G for

details). Since no information on the date of bud dormancy break is available for the calibration, a first structural negative correlation exists between C_{crit} and F_{crit} : the same leaf unfolding date can be obtained with either a long dormancy phase and short bud growth phase or a short dormancy phase and a long bud growth phase. Other structural correlations exist between T_b and C_{crit} on the one hand and d_T/T_{50} and F_{crit} on the other hand: for example, a rapid accumulation of chilling units with a high critical chilling requirement could yield identical results as a slow accumulation with a low critical chilling requirement (i.e. the threshold temperature T_b and the critical chilling requirement C_{crit} are dependent, see Figure G.1.A. in Appendix G).

Consequently, several parameter sets may be statistically equivalent and parameters non-identifiable. In fact, calibration repetitions gave diverging parameter values (Figure 4) while being efficient in distinguishing between species presence and absence (i.e. AUC ~ 0.9 , Figure 3). Thus, even if the calibrated model describes the observed species distribution very well, it does not necessarily mean that parameter values are ecologically relevant. This concern is similar to the criticisms against correlative SDMs, in which parameter values and correlations that well reproduce species ranges not necessarily describe a complex biological reality. In our case, the constraints imposed by the explicit mathematical equations embedded in our models were not sufficient to ensure calibration convergence towards similar solutions that would have suggested a high biological realism. However, it is worth noting that we deliberately chose large parameter ranges (although biologically realistic, i.e. corresponding to the observations made on the different processes modelled across different species) in order to give free rein to the optimization algorithm. As our goal was to assess the performance of CMA-ES objectively, we did not attempt to minimize this non-identifiability issue by restricting the parameter space.

3.3. Methodological issues and perspectives of our study

Our goal here was to investigate the performance of CMA-ES to calibrate quite complex process-based species distribution models using species occurrence data. We did not attempt to validate our parametrizations using independent data, and AUC was only used to determine if model outputs were consistent with species distributions. However, AUC is scale-invariant (it measures how well predictions are ranked rather than their absolute values), and prevents us from obtaining calibrated outputs with more ecological meaning. For example, in PHENOFIT, one species with a fitness of 0.8 could be considered as absent whereas another one as present. Further work could thus be conducted to examine the effects of choosing a different

377 objective function.

378 We compared CMA-ES with a fairly simple but frequently used ABC algorithm because it is widely rec-
379 ognized as an efficient method, and has several advantages including its relative methodological simplicity
380 (which makes it quite easy to implement) and its ability to quantify the uncertainty of the parameters. We
381 did so to provide elements of comparison in order to evaluate CMA-ES performance, and we did not aim
382 to argue for the merits of one approach over another. Higher computational efficiency might be obtained
383 using other types of ABC algorithms, such as Sequential Monte Carlo (Beaumont 2010).

384 It would also be valuable to use a significantly higher computing power, with an adapted version of
385 CMA-ES. To improve the global search performance of CMA-ES, we slightly increased the population size
386 λ (Hansen & Kern 2004) and used a computing cluster to evaluate λ functions in parallel. We were able
387 to use between 40 and 120 cores, which is far from the computing power of some GPUs (> 2000 cores). In
388 this case, choosing a very large population size might not be the best choice. To use efficiently this large
389 parallel computing power, one could rather use a CMA-ES restart strategy (e.g. IPOP-CMA-ES, Auger &
390 Hansen 2005), where population size is successively increased (by a factor of 2), and run these calibrations
391 in parallel. Moreover, when a model requires a high computation time and thus only a small budget can
392 be afforded, the original fitness function could be approximated with a surrogate model in order to reduce
393 the number of original function evaluations required (e.g. Auger *et al.* 2004 ; Loshchilov *et al.* 2013).

394 Finally, several authors advocate for process-based modeling approaches relying upon species response
395 functions that are a priori defined (e.g. Higgins *et al.* 2020). However, the main limitation of such models
396 is the data availability to infer their parameters (Urban *et al.* 2016). Forward parametrization is often long
397 and arduous, and we demonstrated that CMA-ES backward parametrization of PBMs can be a powerful
398 technique. One possible way to facilitate parameter values estimation would be to use backward calibra-
399 tion. In particular, CMA-ES driven by species occurrence data could be used to calibrate submodels whose
400 parameter values can hardly be experimentally measured. However, when a structural correlation exists
401 (as in the leaf unfolding submodel), backward calibration might not provide the right parameter estimates.
402 In such a case, observations and measurements are necessary to determine *a posteriori* which estimates are
403 the most realistic. A combination of both forward and backward calibrations might offer a new perspective
404 for spreading the use of process-based models in predictive ecology, especially for climate change impact
405 studies.

406 **Acknowledgements**

407 The authors would like to thank Hendrik Davi for helping us in using the CASTANEA model. We are also
408 deeply grateful for many helpful comments from Florence Tauc. Finally, we would like to thank François
409 de Coligny, manager of the CAPSIS platform, Gilles Le Moguedec, and the GenOuest and TGCC teams for
410 their support. V.V. was supported by a GAIA doctoral school PhD Fellowship.

411 **Conflict of interest**

412 The authors have no conflicts of interest to declare.

413 **Author contributions**

414 I.C. devised the main conceptual ideas. V.V. worked out the technical details, performed the numerical
415 calculations and wrote the first draft of the manuscript. The two authors discussed the analyses and the
416 results, and contributed to the final manuscript.

417 **Data availability**

418 ERA5-Land dataset is available on the [Copernicus Climate Change Service website](#). EU-SoilHydroGrids
419 is available on the [European Soil Data Centre website](#). SoilGrids250m is available on the [International](#)
420 [Soil Reference and Information Centre website](#). EU-Forest database is available on [FigShare](#). The R code
421 associated with this work will be made available on GitLab upon acceptance of the manuscript.

422 **References**

- 423 Araújo, M.B., Anderson, R.P., Márcia Barbosa, A., Beale, C.M., Dormann, C.F., Early, R., Garcia, R.A.,
424 Guisan, A., Maiorano, L., Naimi, B., O'Hara, R.B., Zimmermann, N.E. & Rahbek, C. (2019). Standards
425 for distribution models in biodiversity assessments. *Science Advances*, **5**. [10.1126/sciadv.aat4858](https://doi.org/10.1126/sciadv.aat4858)
- 426 Auger, A. & Hansen, N. (2005). A restart CMA evolution strategy with increasing population size. *2005*
427 *IEEE Congress on Evolutionary Computation*, pp. 1769–1776 Vol. 2.

- 428 Auger, A., Schoenauer, M. & Vanhaecke, N. (2004). LS-CMA-ES: A Second-Order Algorithm for Covari-
429 ance Matrix Adaptation. *Parallel Problem Solving from Nature - PPSN VIII* (eds X. Yao, E.K. Burke, J.A.
430 Lozano, J. Smith, J.J. Merelo-Guervós, J.A. Bullinaria, J.E. Rowe, P. Tino, A. Kabán & H.-P. Schwefel),
431 pp. 182–191. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg.
- 432 Bäck, T., Foussette, C. & Krause, P. (2013). Empirical Analysis. (eds T. Bäck, C. Foussette & P. Krause),
433 pp. 55–83. Natural Computing Series. Springer, Berlin, Heidelberg.
- 434 Ballings, M. & Van den Poel, D. (2013). *AUC: Threshold independent performance measures for proba-*
435 *bilistic classifiers.*
- 436 Barnosky, A.D., Hadly, E.A., Bascompte, J., Berlow, E.L., Brown, J.H., Fortelius, M., Getz, W.M., Harte,
437 J., Hastings, A., Marquet, P.A., Martinez, N.D., Mooers, A., Roopnarine, P., Vermeij, G., Williams,
438 J.W., Gillespie, R., Kitzes, J., Marshall, C., Matzke, N., Mindell, D.P., Revilla, E. & Smith, A.B. (2012).
439 Approaching a state shift in Earth's biosphere. *Nature*, **486**, 52–58. [10.1038/nature11018](https://doi.org/10.1038/nature11018)
- 440 Beaumont, M.A. (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annual Review*
441 *of Ecology, Evolution, and Systematics*, **41**, 379–406.
- 442 Bengtsson, H. (2021). A unifying framework for parallel and distributed processing in r using futures.
- 443 Bohn, U., Neuhäusl, R., Gisela Gollub, Hettwer, C., Neuhäuslová, Z., Raus, T., Schlüter, H. & Weber, H.
444 (2003). *Map of the natural vegetation of europe - scale 1:2500000*.
- 445 Cai, J. (2019). *Humidity: Calculate water vapor measures from temperature and dew point.*
- 446 Chuine, I. (2000). A Unified Model for Budburst of Trees. *Journal of Theoretical Biology*, **207**, 337–347.
447 [10.1006/jtbi.2000.2178](https://doi.org/10.1006/jtbi.2000.2178)
- 448 Chuine, I. & Beaubien, E.G. (2001). Phenology is a major determinant of tree species range. *Ecology*
449 *Letters*, **4**, 500–510. [10.1046/j.1461-0248.2001.00261.x](https://doi.org/10.1046/j.1461-0248.2001.00261.x)
- 450 Collange, G., Reynaud, S. & Hansen, N. (2010). Covariance matrix adaptation evolution strategy for
451 multidisciplinary optimization of expendable launcher family. *13th AIAA/ISSMO multidisciplinary*
452 *analysis optimization conference.*

- 453 Connolly, S.R., Keith, S.A., Colwell, R.K. & Rahbek, C. (2017). Process, Mechanism, and Modeling in
454 Macroecology. *Trends in Ecology & Evolution*, **32**, 835–844. [10.1016/j.tree.2017.08.011](https://doi.org/10.1016/j.tree.2017.08.011)
- 455 Csilléry, K., Blum, M.G.B., Gaggiotti, O.E. & François, O. (2010). Approximate Bayesian Computation
456 (ABC) in practice. *Trends in Ecology & Evolution*, **25**, 410–418. [10.1016/j.tree.2010.04.001](https://doi.org/10.1016/j.tree.2010.04.001)
- 457 Davi, H. & Cailleret, M. (2017). Assessing drought-driven mortality trees with physiological process-
458 based models. *Agricultural and Forest Meteorology*, **232**, 279–290. [10.1016/j.agrformet.2016.08.019](https://doi.org/10.1016/j.agrformet.2016.08.019)
- 459 Davi, H., Dufrêne, E., Francois, C., Le Maire, G., Loustau, D., Bosc, A., Rambal, S., Granier, A. & Moors,
460 E. (2006). Sensitivity of water and carbon fluxes to climate changes from 1960 to 2100 in European
461 forest ecosystems. *Agricultural and Forest Meteorology*, **141**, 35–56. [10.1016/j.agrformet.2006.09.003](https://doi.org/10.1016/j.agrformet.2006.09.003)
- 462 Davi, H., Dufrêne, E., Granier, A., Le Dantec, V., Barbaroux, C., François, C. & Bréda, N. (2005). Mod-
463 elling carbon and water cycles in a beech forest: Part II.: Validation of the main processes from
464 organ to stand scale. *Ecological Modelling*, **185**, 387–405. [10.1016/j.ecolmodel.2005.01.003](https://doi.org/10.1016/j.ecolmodel.2005.01.003)
- 465 Delpierre, N., Soudani, K., François, C., Le Maire, G., Bernhofer, C., Kutsch, W., Misson, L., Rambal, S.,
466 Vesala, T. & Dufrêne, E. (2012). Quantifying the influence of climate and biological drivers on the
467 interannual variability of carbon exchanges in European forests through process-based modelling.
468 *Agricultural and Forest Meteorology*, **154-155**, 99–112. [10.1016/j.agrformet.2011.10.010](https://doi.org/10.1016/j.agrformet.2011.10.010)
- 469 Dormann, C.F., Schymanski, S.J., Cabral, J., Chuine, I., Graham, C., Hartig, F., Kearney, M., Morin,
470 X., Römermann, C., Schröder, B. & Singer, A. (2012). Correlation and process in species distri-
471 bution models: Bridging a dichotomy. *Journal of Biogeography*, **39**, 2119–2131. [2699.2011.02659.x](https://doi.org/10.1111/j.1365-
472 2699.2011.02659.x)
- 473 Dufrêne, E., Davi, H., François, C., Maire, G. le, Dantec, V.L. & Granier, A. (2005). Modelling carbon and
474 water cycles in a beech forest: Part I: Model description and uncertainty analysis on modelled NEE.
475 *Ecological Modelling*, **185**, 407–436. [10.1016/j.ecolmodel.2005.01.004](https://doi.org/10.1016/j.ecolmodel.2005.01.004)
- 476 Duputié, A., Rutschmann, A., Ronce, O. & Chuine, I. (2015). Phenological plasticity will not help all
477 species adapt to climate change. *Global Change Biology*, **21**, 3062–3073. [10.1111/gcb.12914](https://doi.org/10.1111/gcb.12914)

- 478 Duputié, A., Zimmermann, N.E. & Chuine, I. (2014). Where are the wild things? Why we need better
479 data on species distribution. *Global Ecology and Biogeography*, **23**, 457–467. [10.1111/geb.12118](https://doi.org/10.1111/geb.12118)
- 480 Evans, M.R. (2012). Modelling ecological systems in a changing world. *Philosophical Transactions of the*
481 *Royal Society B: Biological Sciences*, **367**, 181–190. [10.1098/rstb.2011.0172](https://doi.org/10.1098/rstb.2011.0172)
- 482 Fitzpatrick, M.C., Blois, J.L., Williams, J.W., Nieto-Lugilde, D., Maguire, K.C. & Lorenz, D.J. (2018). How
483 will climate novelty influence ecological forecasts? Using the Quaternary to assess future reliability.
484 *Global Change Biology*, **24**, 3575–3586. [10.1111/gcb.14138](https://doi.org/10.1111/gcb.14138)
- 485 Fourcade, Y., Besnard, A.G. & Secondi, J. (2018). Paintings predict the distribution of species, or the
486 challenge of selecting environmental predictors and evaluation statistics. *Global Ecology and Bio-*
487 *geography*, **27**, 245–256. [10.1111/geb.12684](https://doi.org/10.1111/geb.12684)
- 488 Gagné, C., Beaulieu, J., Parizeau, M. & Thibault, S. (2008). Human-competitive lens system design with
489 evolution strategies. *Applied Soft Computing*, **8**, 1439–1452. [10.1016/j.asoc.2007.10.018](https://doi.org/10.1016/j.asoc.2007.10.018)
- 490 Gardner, E., Breeze, T.D., Clough, Y., Smith, H.G., Baldock, K.C.R., Campbell, A., Garratt, M.P.D., Gille-
491 spie, M.A.K., Kunin, W.E., McKerchar, M., Memmott, J., Potts, S.G., Senapathi, D., Stone, G.N., Wäck-
492 ers, F., Westbury, D.B., Wilby, A. & Oliver, T.H. (2020). Reliably predicting pollinator abundance:
493 Challenges of calibrating process-based ecological models. *Methods in Ecology and Evolution*, **11**,
494 1673–1689. [10.1111/2041-210X.13483](https://doi.org/10.1111/2041-210X.13483)
- 495 Gauzere, J., Teuf, B., Davi, H., Chevin, L.-M., Caillard, T., Leys, B., Delzon, S., Ronce, O. & Chuine,
496 I. (2020). Where is the optimum? Predicting the variation of selection along climatic gradients
497 and the adaptive value of plasticity. A case study on tree phenology. *Evolution Letters*, **4**, 109–123.
498 [10.1002/evl3.160](https://doi.org/10.1002/evl3.160)
- 499 GBIF. (2022). The global biodiversity information facility. <https://www.gbif.org> [accessed 24 January
500 2022]
- 501 Guisan, A. & Thuiller, W. (2005). Predicting species distribution: Offering more than simple habitat
502 models. *Ecology Letters*, **8**, 993–1009. [10.1111/j.1461-0248.2005.00792.x](https://doi.org/10.1111/j.1461-0248.2005.00792.x)

- 503 Guo, D., Westra, S. & Maier, H.R. (2016). An R package for modelling actual, potential and reference
504 evapotranspiration. *Environmental Modelling & Software*, **78**, 216–224. [10.1016/j.envsoft.2015.12.019](https://doi.org/10.1016/j.envsoft.2015.12.019)
- 505 Hansen, N. (2016). The CMA evolution strategy: A tutorial.
- 506 Hansen, N. (2006). The CMA Evolution Strategy: A Comparing Review. (eds J.A. Lozano, P. Larrañaga,
507 I. Inza & E. Bengoetxea), pp. 75–102. Studies in Fuzziness and Soft Computing. Springer, Berlin,
508 Heidelberg.
- 509 Hansen, N., Auger, A., Ros, R., Finck, S. & Posik, P. (2010). Comparing Results of 31 Algorithms from the
510 Black-Box Optimization Benchmarking BBOB-2009. *ACM-GECCO Genetic and Evolutionary Com-*
511 *putation Conference*. Portland, United States.
- 512 Hansen, N. & Kern, S. (2004). Evaluating the CMA Evolution Strategy on Multimodal Test Functions.
513 *Parallel Problem Solving from Nature - PPSN VIII* (eds X. Yao, E.K. Burke, J.A. Lozano, J. Smith,
514 J.J. Merelo-Guervós, J.A. Bullinaria, J.E. Rowe, P. Tino, A. Kabán & H.-P. Schwefel), pp. 282–291.
515 Lecture Notes in Computer Science. Springer, Berlin, Heidelberg.
- 516 Hansen, N., Niederberger, A.S.P., Guzzella, L. & Koumoutsakos, P. (2009). A Method for Handling
517 Uncertainty in Evolutionary Optimization With an Application to Feedback Control of Combustion.
518 *IEEE Transactions on Evolutionary Computation*, **13**, 180–197. [10.1109/TEVC.2008.924423](https://doi.org/10.1109/TEVC.2008.924423)
- 519 Hansen, N. & Ostermeier, A. (1996). Adapting arbitrary normal mutation distributions in evolution
520 strategies: The covariance matrix adaptation. *Proceedings of IEEE International Conference on Evo-*
521 *lutionary Computation*, pp. 312–317.
- 522 Hansen, N. & Ostermeier, A. (2001). Completely Derandomized Self-Adaptation in Evolution Strategies.
523 *Evolutionary Computation*, **9**, 159–195. [10.1162/106365601750190398](https://doi.org/10.1162/106365601750190398)
- 524 Hansen, N., Ros, R., Mauny, N., Schoenauer, M. & Auger, A. (2011). Impacts of invariance in search:
525 When CMA-ES and PSO face ill-conditioned and non-separable problems. *Applied Soft Computing*,
526 **11**, 5755–5769. [10.1016/j.asoc.2011.03.001](https://doi.org/10.1016/j.asoc.2011.03.001)
- 527 Hartig, F., Dislich, C., Wiegand, T. & Huth, A. (2014). Technical Note: Approximate Bayesian param-
528 eterization of a process-based tropical forest model. *Biogeosciences*, **11**, 1261–1272. [10.5194/bg-11-](https://doi.org/10.5194/bg-11-)

- 530 Hengl, T., Jesus, J.M. de, Heuvelink, G.B.M., Gonzalez, M.R., Kilibarda, M., Blagotic, A., Shangguan, W.,
 531 Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Bat-
 532 jes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S. & Kempen, B. (2017). SoilGrids250m:
 533 Global gridded soil information based on machine learning. *PLoS ONE*, **12**, e0169748. [10.1371/journal.pone.0169748](https://doi.org/10.1371/journal.pone.0169748)
- 535 Higgins, S.I., Larcombe, M.J., Beeton, N.J., Conradi, T. & Nottebrock, H. (2020). Predictive ability of a
 536 process-based versus a correlative species distribution model. *Ecology and Evolution*, **10**, 11043–
 537 11054. [10.1002/ece3.6712](https://doi.org/10.1002/ece3.6712)
- 538 Higgins, S.I., O'Hara, R.B. & Römermann, C. (2012). A niche for biology in species distribution models.
 539 *Journal of Biogeography*, **39**, 2091–2095. [10.1111/jbi.12029](https://doi.org/10.1111/jbi.12029)
- 540 Hijmans, R.J., Phillips, S., Leathwick, J. & Elith, J. (2021). *Dismo: Species distribution modeling*.
- 541 Hill, A., Laneurit, J., Lenain, R. & Lucet, E. (2020). Online gain setting method for path tracking us-
 542 ing CMA-ES: Application to off-road mobile robot control. *IROS 2020, International Conference on*
 543 *Intelligent Robots and Systems*. Las Vegas, United States.
- 544 Jalas, J. & Suominen, J. (1972–2005). *Atlas florae europaea*. Committee for Mapping the Flora of Europe;
 545 Societas Biologica Fennica Vanamo, Helsinki, Finland.
- 546 Journé, V., Barnagaud, J., Bernard, C., Crochet, P. & Morin, X. (2020). Correlative climatic niche models
 547 predict real and virtual species distributions equally well. *Ecology*, **101**. [10.1002/ecy.2912](https://doi.org/10.1002/ecy.2912)
- 548 Kleidon, A. & Mooney, H.A. (2000). A global distribution of biodiversity inferred from climatic con-
 549 straints: Results from a process-based modelling study. *Global Change Biology*, **6**, 507–523. [10.1046/j.1365-2486.2000.00332.x](https://doi.org/10.1046/j.1365-2486.2000.00332.x)
- 551 Lagarrigues, G., Jabot, F., Lafond, V. & Courbaud, B. (2015). Approximate Bayesian computation to re-
 552 calibrate individual-based models with population data: Illustration with a forest simulation model.
 553 *Ecological Modelling*, **306**, 278–286. [10.1016/j.ecolmodel.2014.09.023](https://doi.org/10.1016/j.ecolmodel.2014.09.023)

- 554 Lenton, T.M., Held, H., Kriegler, E., Hall, J.W., Lucht, W., Rahmstorf, S. & Schellnhuber, H.J. (2008).
555 Tipping elements in the Earth's climate system. *Proceedings of the National Academy of Sciences*,
556 **105**, 1786–1793. [10.1073/pnas.0705414105](https://doi.org/10.1073/pnas.0705414105)
- 557 Leroy, B., Delsol, R., Hugueny, B., Meynard, C.N., Barhoumi, C., Barbet-Massin, M. & Bellard, C. (2018).
558 Without quality presence-absence data, discrimination metrics such as TSS can be misleading mea-
559 sures of model performance. *Journal of Biogeography*, **45**, 1994–2002. [10.1111/jbi.13402](https://doi.org/10.1111/jbi.13402)
- 560 Lobo, J.M., Jiménez-Valverde, A. & Real, R. (2008). AUC: A misleading measure of the performance
561 of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151. [10.1111/j.1466-8238.2007.00358.x](https://doi.org/10.1111/j.1466-8238.2007.00358.x)
- 563 Loshchilov, I., Schoenauer, M. & Sèbag, M. (2013). Bi-population CMA-ES algorithms with surrogate
564 models and line searches. *Proceedings of the 15th annual conference companion on Genetic and evolu-*
565 *tionary computation*, pp. 1177–1184. GECCO '13 Companion. Association for Computing Machin-
566 ery, New York, NY, USA.
- 567 Mauri, A., Strona, G. & San-Miguel-Ayanz, J. (2017). EU-Forest, a high-resolution tree occurrence
568 dataset for Europe. *Scientific Data*, **4**, 160123. [10.1038/sdata.2016.123](https://doi.org/10.1038/sdata.2016.123)
- 569 Merow, C., Smith, M.J., Edwards, T.C., Guisan, A., McMahon, S.M., Normand, S., Thuiller, W., Wüest,
570 R.O., Zimmermann, N.E. & Elith, J. (2014). What do we gain from simplicity versus complexity in
571 species distribution models? *Ecography*, **37**, 1267–1281. [10.1111/ecog.00845](https://doi.org/10.1111/ecog.00845)
- 572 Monnet, A.-C., Cilleros, K., Médail, F., Albassatneh, M.C., Arroyo, J., Bacchetta, G., Bagnoli, F., Ba-
573 rrina, Z., Cartereau, M., Casajus, N., Dimopoulos, P., Domina, G., Doxa, A., Escudero, M., Fady, B.,
574 Hampe, A., Matevski, V., Misfud, S., Nikolic, T., Pavon, D., Roig, A., Bareja, E.S., Spanu, I., Strid,
575 A., Vendramin, G.G. & Leriche, A. (2021). WOODIV, a database of occurrences, functional traits,
576 and phylogenetic data for all Euro-Mediterranean trees. *Scientific Data*, **8**, 89. [10.1038/s41597-021-00873-3](https://doi.org/10.1038/s41597-021-00873-3)
- 578 Morin, X., Augspurger, C. & Chuine, I. (2007). Process-Based Modeling of Species' Distributions: What
579 Limits Temperate Tree Species' Range Boundaries? *Ecology*, **88**, 2280–2291. [10.1890/06-1591.1](https://doi.org/10.1890/06-1591.1)

- 580 Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., Faure, D., Garnier, E.,
581 Gimenez, O., Huneman, P., Jabot, F., Jarne, P., Joly, D., Julliard, R., Kéfi, S., Kergoat, G.J., Lavorel,
582 S., Le Gall, L., Meslin, L., Morand, S., Morin, X., Morlon, H., Pinay, G., Pradel, R., Schurr, F.M.,
583 Thuiller, W. & Loreau, M. (2015). REVIEW: Predictive ecology in a changing world. *Journal of*
584 *Applied Ecology*, **52**, 1293–1310. [10.1111/1365-2664.12482](https://doi.org/10.1111/1365-2664.12482)
- 585 Muñoz Sabater, J. (2021). ERA5-land hourly data from 1950 to 1980. [https://cds.climate.copernicus.eu/](https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land)
586 [cdsapp#!/dataset/reanalysis-era5-land](https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land)
- 587 Muñoz Sabater, J. (2019). ERA5-land hourly data from 1981 to present. [https://cds.climate.copernicus.](https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land)
588 [eu/cdsapp#!/dataset/reanalysis-era5-land](https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land)
- 589 Pearman, P.B., Guisan, A., Broennimann, O. & Randin, C.F. (2008). Niche dynamics in space and time.
590 *Trends in Ecology & Evolution*, **23**, 149–158. [10.1016/j.tree.2007.11.005](https://doi.org/10.1016/j.tree.2007.11.005)
- 591 Radeloff, V.C., Williams, J.W., Bateman, B.L., Burke, K.D., Carter, S.K., Childress, E.S., Cromwell, K.J.,
592 Gratton, C., Hasley, A.O., Kraemer, B.M., Latzka, A.W., Marin-Spiotta, E., Meine, C.D., Munoz, S.E.,
593 Neeson, T.M., Pidgeon, A.M., Rissman, A.R., Rivera, R.J., Szymanski, L.M. & Usinowicz, J. (2015).
594 The rise of novelty in ecosystems. *Ecological Applications*, **25**, 2051–2068. [10.1890/14-1781.1](https://doi.org/10.1890/14-1781.1)
- 595 Saltré, F., Saint-Amant, R., Gritti, E.S., Brewer, S., Gaucherel, C., Davis, B.A.S. & Chuine, I. (2013).
596 Climate or migration: What limited European beech post-glacial colonization? *Global Ecology and*
597 *Biogeography*, **22**, 1217–1227. [10.1111/geb.12085](https://doi.org/10.1111/geb.12085)
- 598 Santini, L., Benítez-López, A., Maiorano, L., Cengic, M. & Huijbregts, M.A.J. (2021). Assessing the
599 reliability of species distribution projections in climate change research. *Diversity and Distributions*,
600 **27**, 1035–1050. [10.1111/ddi.13252](https://doi.org/10.1111/ddi.13252)
- 601 Singer, A., Johst, K., Banitz, T., Fowler, M.S., Groeneveld, J., Gutiérrez, A.G., Hartig, F., Krug, R.M., Liess,
602 M., Matlack, G., Meyer, K.M., Pe'er, G., Radchuk, V., Voinopol-Sassu, A.-J. & Travis, J.M.J. (2016).
603 Community dynamics under environmental change: How can next generation mechanistic models
604 improve projections of species distributions? *Ecological Modelling*, **326**, 63–74. [10.1016/j.ecolmodel.2015.11.007](https://doi.org/10.1016/j.ecolmodel.2015.11.007)
- 605 Steffen, W., Rockström, J., Richardson, K., Lenton, T.M., Folke, C., Liverman, D., Summerhayes, C.P.,
606 Barnosky, A.D., Cornell, S.E., Crucifix, M., Donges, J.F., Fetzer, I., Lade, S.J., Scheffer, M., Winkel-

- 607 mann, R. & Schellnhuber, H.J. (2018). Trajectories of the Earth System in the Anthropocene. *Proceedings of the National Academy of Sciences*, **115**, 8252–8259. [10.1073/pnas.1810141115](https://doi.org/10.1073/pnas.1810141115)
- 608
- 609 Tóth, B., Weynants, M., Pásztor, L. & Hengl, T. (2017). 3D soil hydraulic database of Europe at 250 m
610 resolution. *Hydrological Processes*, **31**, 2662–2666. [10.1002/hyp.11203](https://doi.org/10.1002/hyp.11203)
- 611
- 612 Trautmann, H., Mersmann, O. & Arnu, D. (2011). *Cmaes: Covariance matrix adapting evolutionary*
613 *strategy*.
- 614
- 615 Urban, M.C. (2015). Accelerating extinction risk from climate change. *Science*, **348**, 571–573. [10.1126/science.aaa4984](https://doi.org/10.1126/science.aaa4984)
- 616
- 617 Urban, M.C., Bocedi, G., Hendry, A.P., Mihoub, J.-B., Pe'er, G., Singer, A., Bridle, J.R., Crozier, L.G.,
618 De Meester, L., Godsoe, W., Gonzalez, A., Hellmann, J.J., Holt, R.D., Huth, A., Johst, K., Krug, C.B.,
Leadley, P.W., Palmer, S.C.F., Pantel, J.H., Schmitz, A., Zollner, P.A. & Travis, J.M.J. (2016). Improving
the forecast for biodiversity under climate change. *Science*, **353**, aad8466. [10.1126/science.aad8466](https://doi.org/10.1126/science.aad8466)
- 619
- 620 Vaart, E. van der, Beaumont, M.A., Johnston, A.S.A. & Sibly, R.M. (2015). Calibration and evaluation
621 of individual-based models using Approximate Bayesian Computation. *Ecological Modelling*, **312**,
182–190. [10.1016/j.ecolmodel.2015.05.020](https://doi.org/10.1016/j.ecolmodel.2015.05.020)
- 622
- 623 Warren, D.L., Dornburg, A., Zapfe, K. & Iglesias, T.L. (2021). The effects of climate change on Australia's
only endemic Pokémon: Measuring bias in species distribution models. *Methods in Ecology and*
624 *Evolution*, **12**, 985–995. [10.1111/2041-210X.13591](https://doi.org/10.1111/2041-210X.13591)
- 625
- 626 Williams, J.W., Jackson, S.T. & Kutzbach, J.E. (2007). Projected distributions of novel and disappearing
climates by 2100 AD. *Proceedings of the National Academy of Sciences*, **104**, 5738–5742. [10.1073/pnas.0606292104](https://doi.org/10.1073/pnas.0606292104)
- 627
- 628 Zurell, D., Thuiller, W., Pagel, J., Cabral, J.S., Münkemüller, T., Gravel, D., Dullinger, S., Normand,
S., Schiffers, K.H., Moore, K.A. & Zimmermann, N.E. (2016). Benchmarking novel approaches for
629 modelling species range dynamics. *Global Change Biology*, **22**, 2651–2664. [10.1111/gcb.13251](https://doi.org/10.1111/gcb.13251)

630 Supplementary Appendix A: Insights on the models

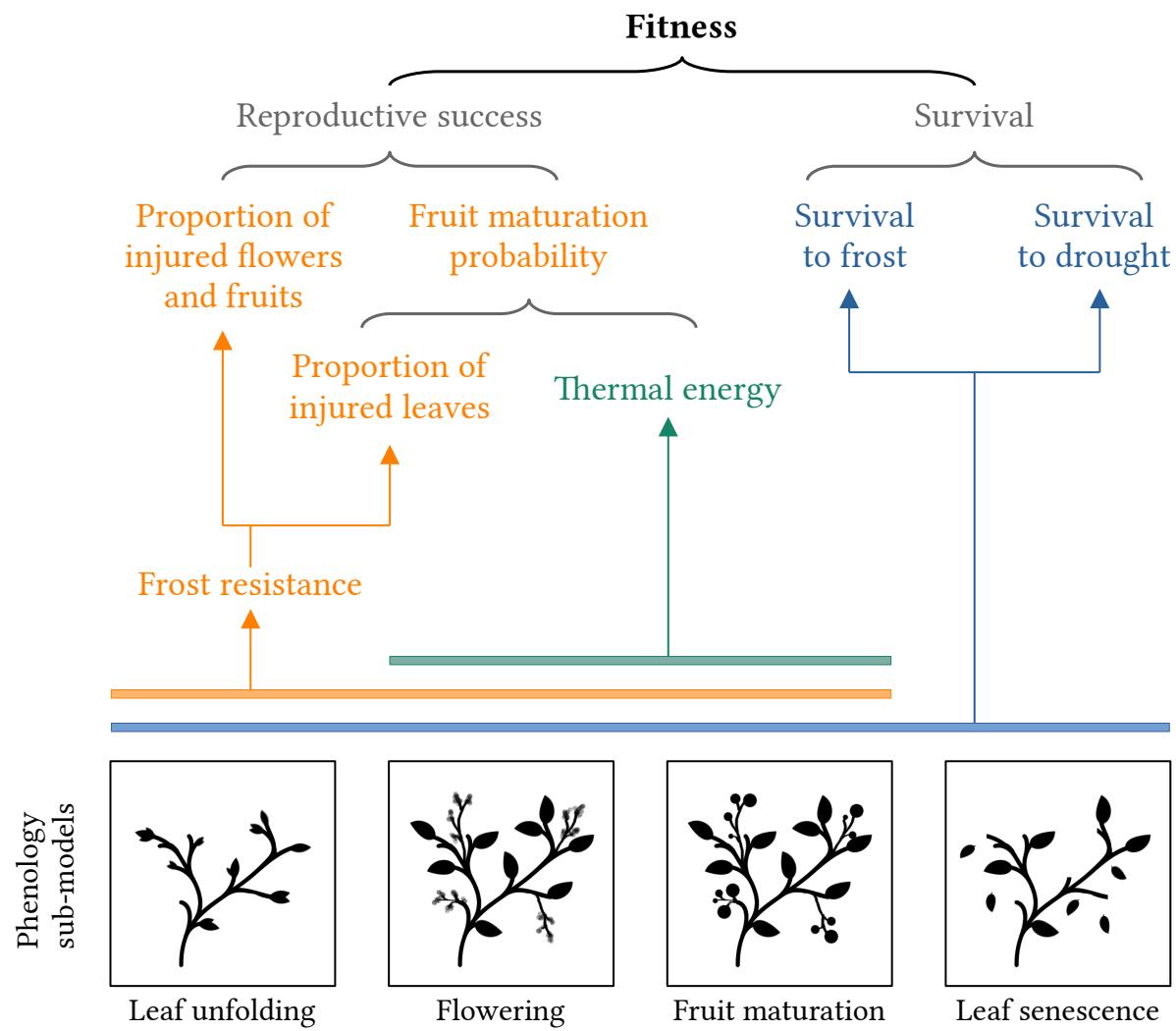


Figure A.1: PHENOFIT model in a nutshell.

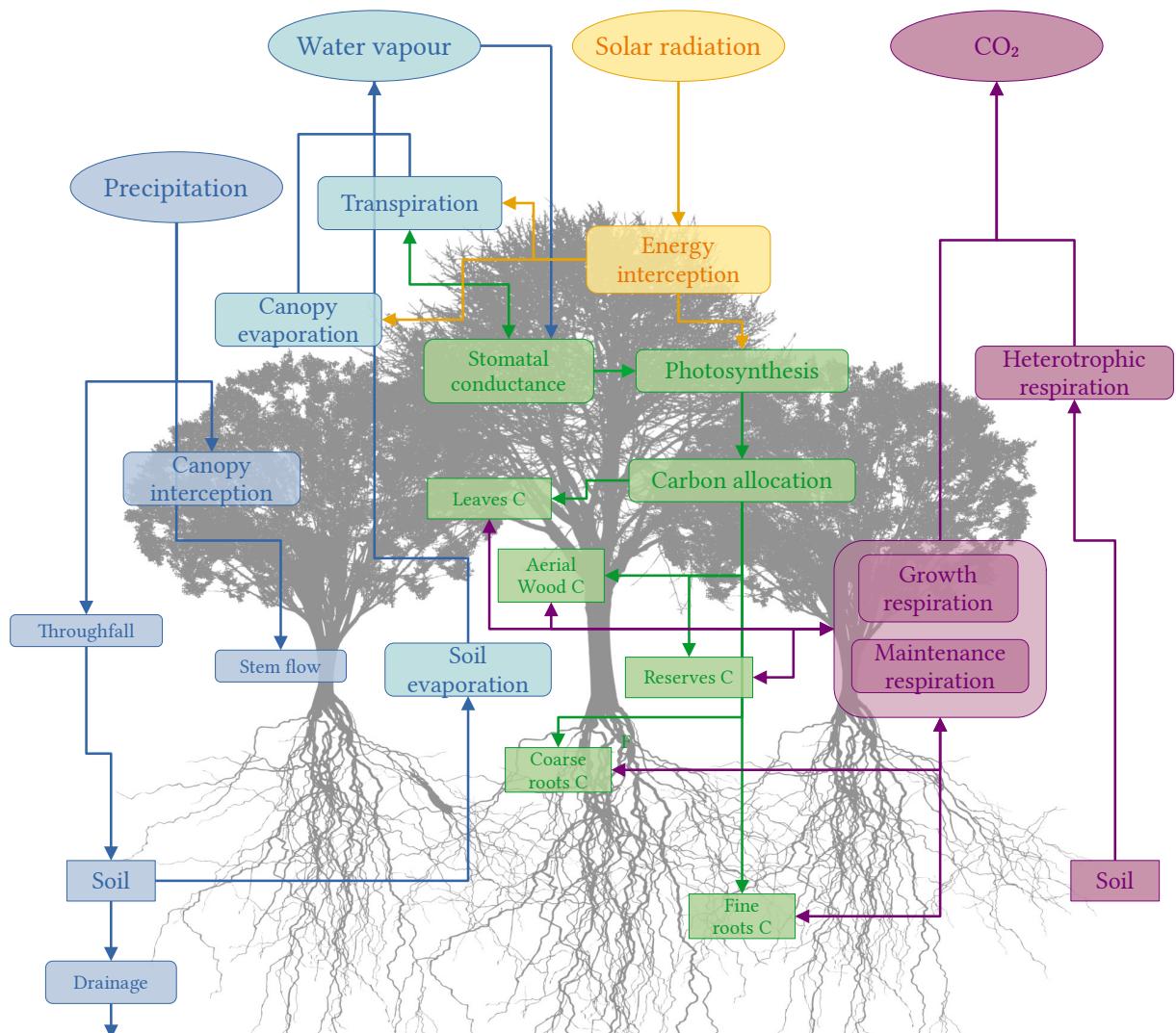


Figure A.2: CASTANEA model in a nutshell.

631 **Supplementary Appendix B: Processing of occurrence data**

632

633

Table B.1: GBIF download links

Species	Number of occurrences	Download link
<i>Fagus sylvatica</i>	718.898	https://doi.org/10.15468/dl.e9wasa
<i>Quercus ilex</i>	78.979	https://doi.org/10.15468/dl.2a4haw
<i>Abies alba</i>	119.891	https://doi.org/10.15468/dl.my6c9t

634

635

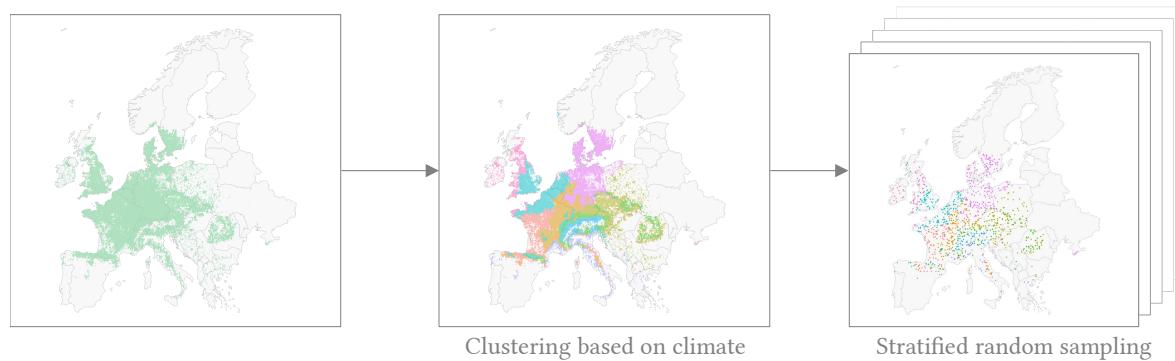


Figure B.1: Stratified random sampling of beech presence records based on climate clusters.

B-2

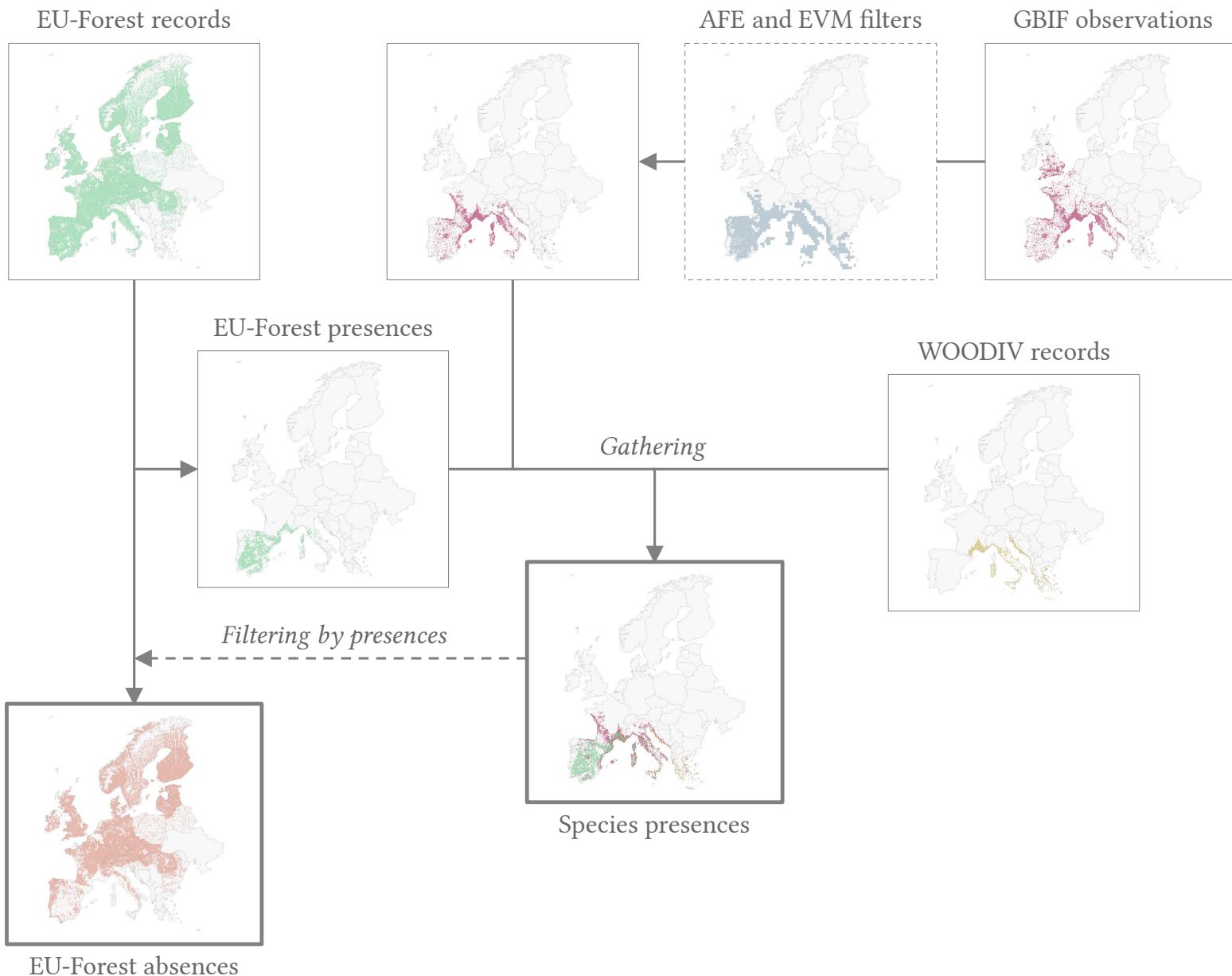


Figure B.2: Processing of holm oak occurrence records. GBIF: Global Biodiversity Information Facility, AFE: Atlas Flora Europeae, EVM: EuroVegMap.

636 **Supplementary Appendix C: Species distributions**

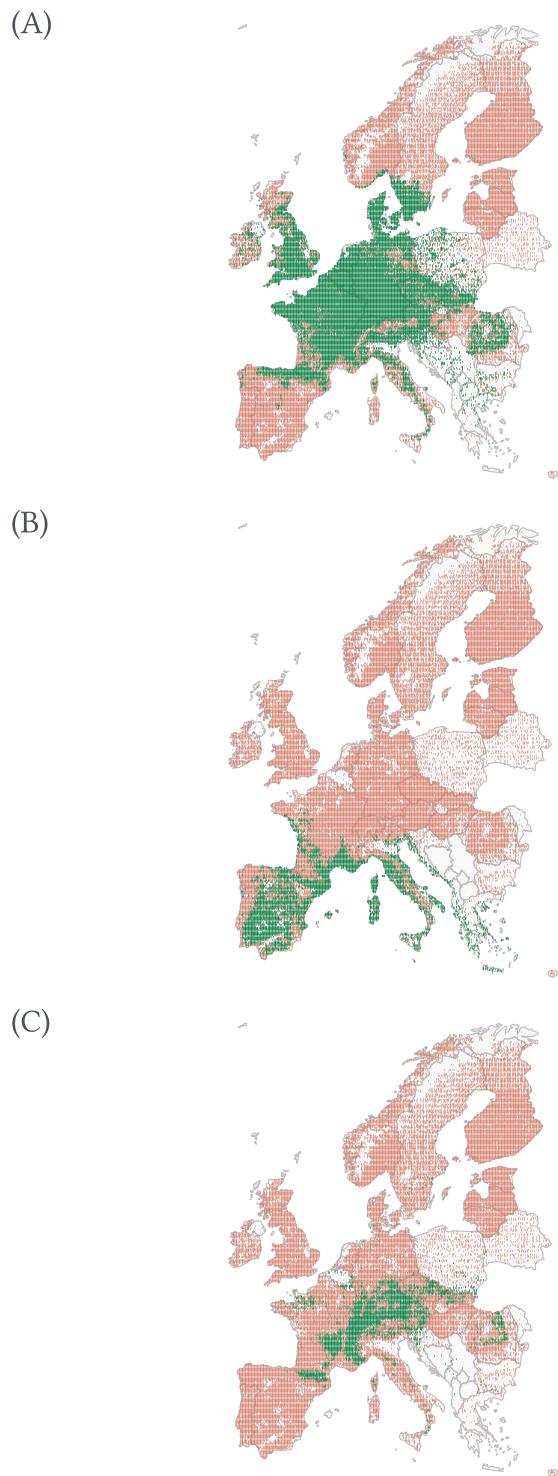


Figure C.1: Species distributions of (A) beech, (B) holm oak and (C) silver fir. Green cells are 0.1° cells where species is present, orange cells where species is supposed to be absent.

637 **Supplementary Appendix D: Handling constraints in CMA-ES**

638

639

640 *Box constraint handling*

641 With this constraint handling - implemented by default in the R package *cmaes* ([Trautmann et al. 2011](#))
642 - each evaluated solution is guaranteed to lie within the feasible space. Let's say we have a parameter
643 vector x . For each parameter x_i , we have a lower bound lb_i and an upper bound ub_i . If a parameter x_i
644 violates one of this bound, we set x_i to a new value $x_i^{repaired}$ equal to the closest boundary value (lb_i or
645 ub_i). We thus obtained a new parameter set $x^{repaired}$, with a minimal $\|x - x^{repaired}\|$ value. This new
646 feasible solution $x^{repaired}$ is used for the evaluation of the objective function $AUC_{model}(x^{repaired})$, and
647 to compute a penalty term $pen = \sum_i (x_i - x_i^{repaired})^2 = \|x^{repaired} - x\|^2$. Then $x^{repaired}$ is discarded,
648 and the algorithm computes the penalized objective function of $x^{repaired}$ as follows: $AUC_{model}(x) =$
649 $AUC_{model}(x^{repaired}) + pen$. This boundary handling could be improved with adaptive weights (see
650 [Hansen et al. 2009](#)).

651 *Ecological infeasibility constraint*

652 We added a simple way to handle ecological constraint (e.g. unfolding before flowering in beech mixed
653 bud) with a death penalty. When a parameter vector x violates a constraint, it is rejected and generated
654 again. The main drawback of this approach is that CMA-ES does not use information from unfeasible
655 points. An other approach could be to set $AUC_{model}(x) = 0$. However, as our feasible space was large,
656 the death penalty constraint worked well in our case.

657 **Supplementary Appendix E: holm oak and silver fir calibrations**

658

659

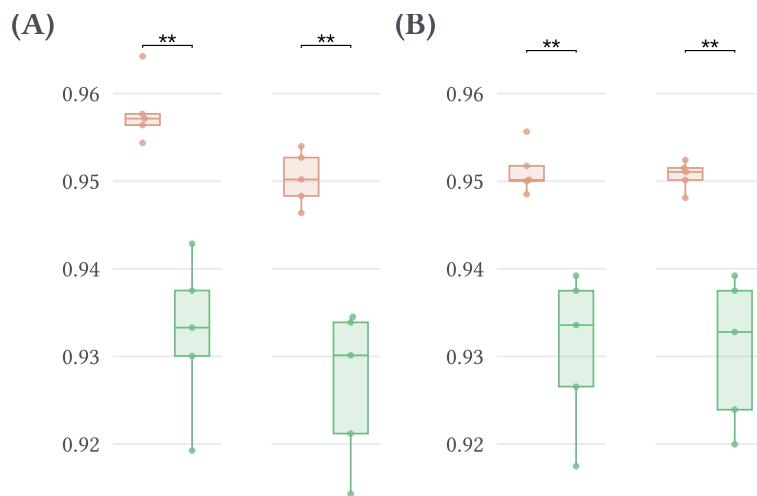


Figure E.1: Comparison of CMA-ES and ABC-rejection methods, with two holm oak occurrence subsets, on (A) calibration AUC (only calibration points) and (B) total AUC (every presence/absence points). Each point is a calibration run. The black horizontal bars represent the pairwise Mann–Whitney tests between the two methods on the same subset. CMA-ES is red and ABC green.

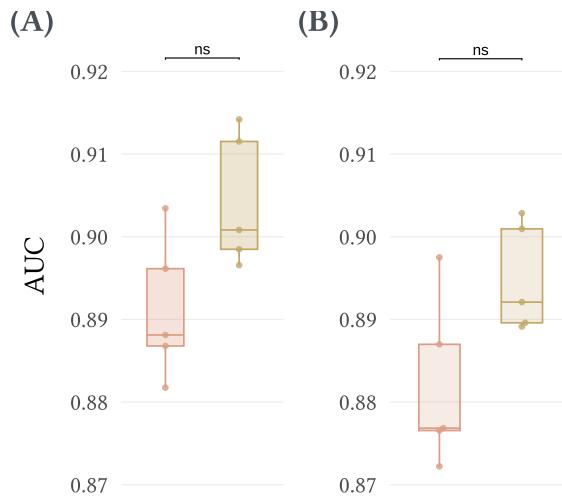


Figure E.2: CMA-ES calibration using the PHENOFIT model and silver fir: (A) calibration AUC (only calibration cells) and (B) total AUC (every presence/absence cells). Each color is a different sub-sampling of occurrence data, each point is a calibration run. The black horizontal bars represent the pairwise Mann–Whitney tests between the two subsets.

660 **Supplementary Appendix F: Raw model outputs**

661

662

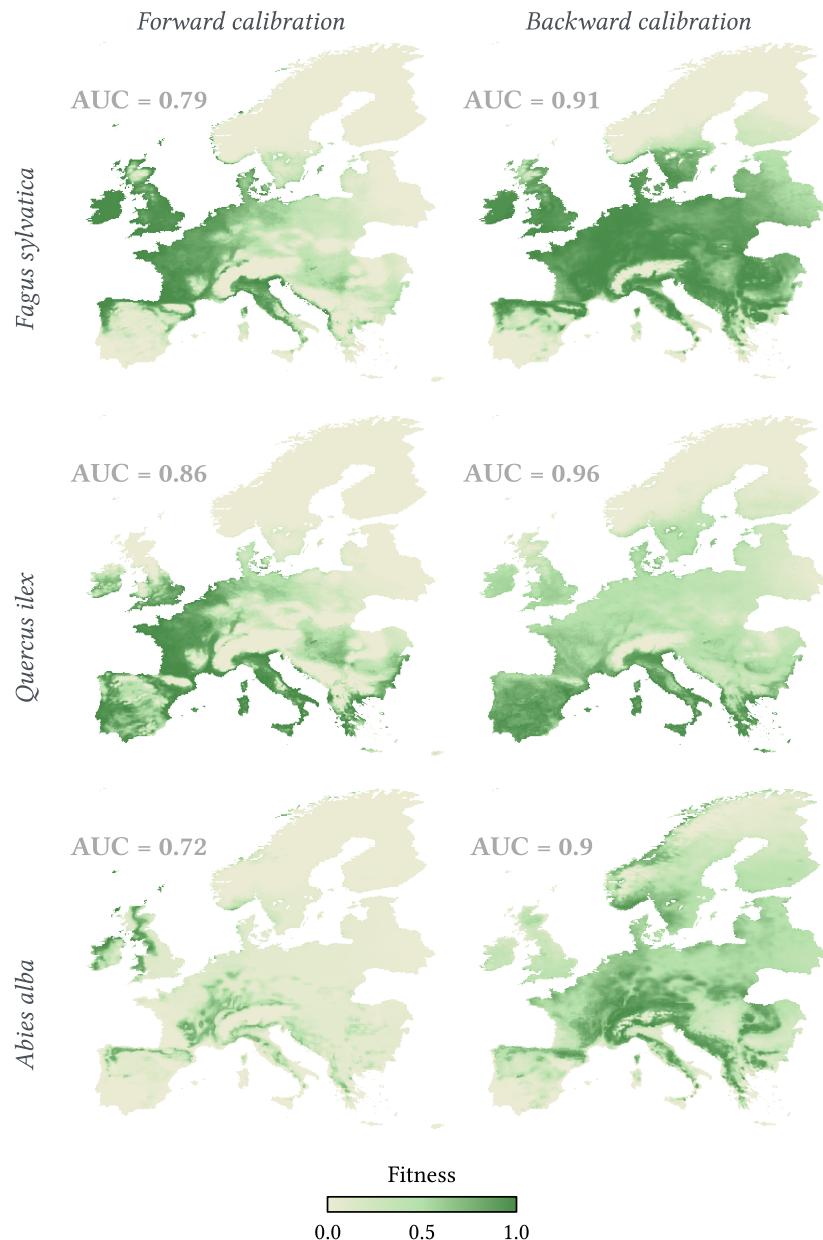


Figure F.1: Fitness index predicted by PHENOFIT with the forward and the backward calibrations.

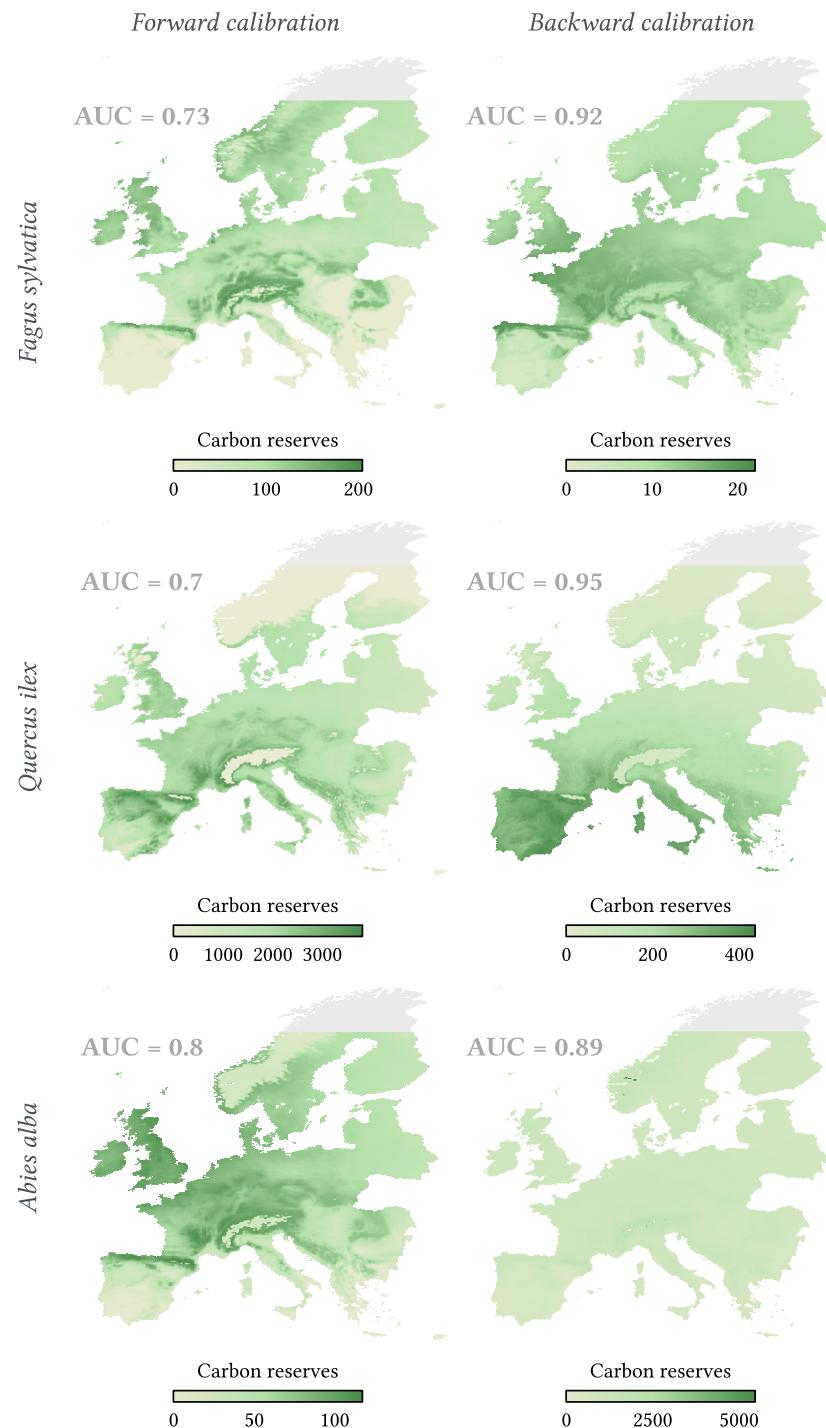


Figure F.2: Carbon reserves predicted by CASTANEA with the forward and the backward calibrations.

663 **Supplementary Appendix G: Leaf unfolding submodel**

664

665

666 *Fagus sylvatica* leaf unfolding submodel

667 This model, called UniChill ([Chuine 2000](#)), is a sequential two-phase model (endodormancy and ecodormancy phases).

669 The endodormancy phase begins at day t_0 . The daily rate of chilling R_c is defined as a threshold function of the daily mean temperature T_d :

$$R_c(T_d) = \begin{cases} 0 & T_d \geq T_b \\ 1 & T_d < T_b \end{cases}$$

671 where T_b is the threshold temperature below which the bud accumulates chilling units.

672 The endodormancy releases at day t_c when the accumulated rate of chilling has reached the level C_{crit} :

$$\sum_{t_0}^{t_c} R_c(T_d) \geq C_{crit}$$

673 Then, the ecodormancy phase begins. The daily rate of forcing R_f is defined as a sigmoid function of
674 the daily mean temperature T_d :

$$R_f(T_d) = \frac{1}{1 + e^{-d_T(T_d - T_{50})}}$$

675 where d_T is the slope and T_{50} the mid-response temperature. Bud break occurs at day t_f when the
676 accumulated rate of forcing has reached the level F_{crit} :

$$\sum_{t_c}^{t_f} R_f(T_d) \geq F_{crit}$$

677 Thus, the UniChill model has 6 parameters: t_0 , T_b and C_{crit} for the first phase, d_T , T_{50} and F_{crit} for the
678 second phase.

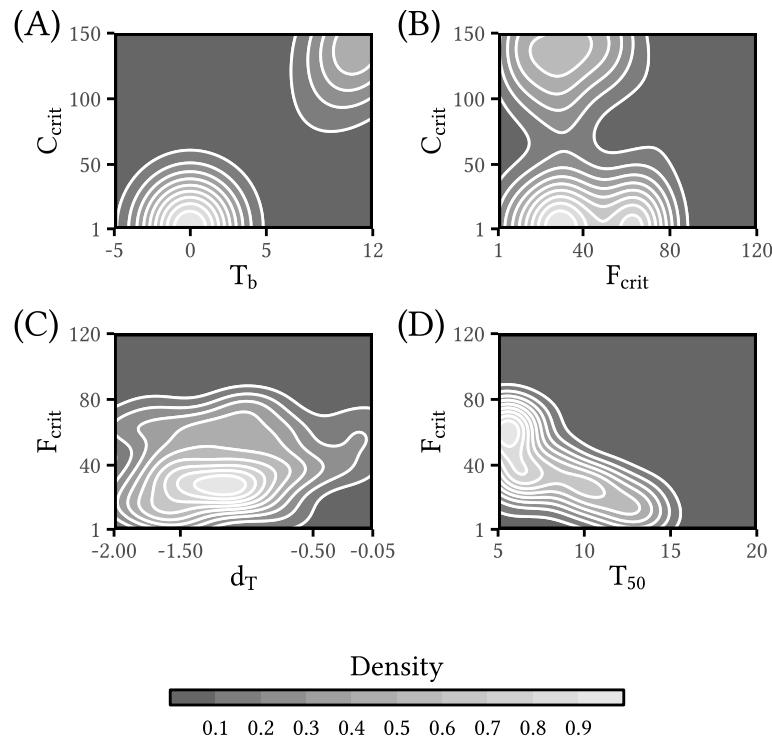


Figure G.1: Beech leaf unfolding model parameter density. Y-axis and X-axis limits are lower and upper bounds used during calibration.

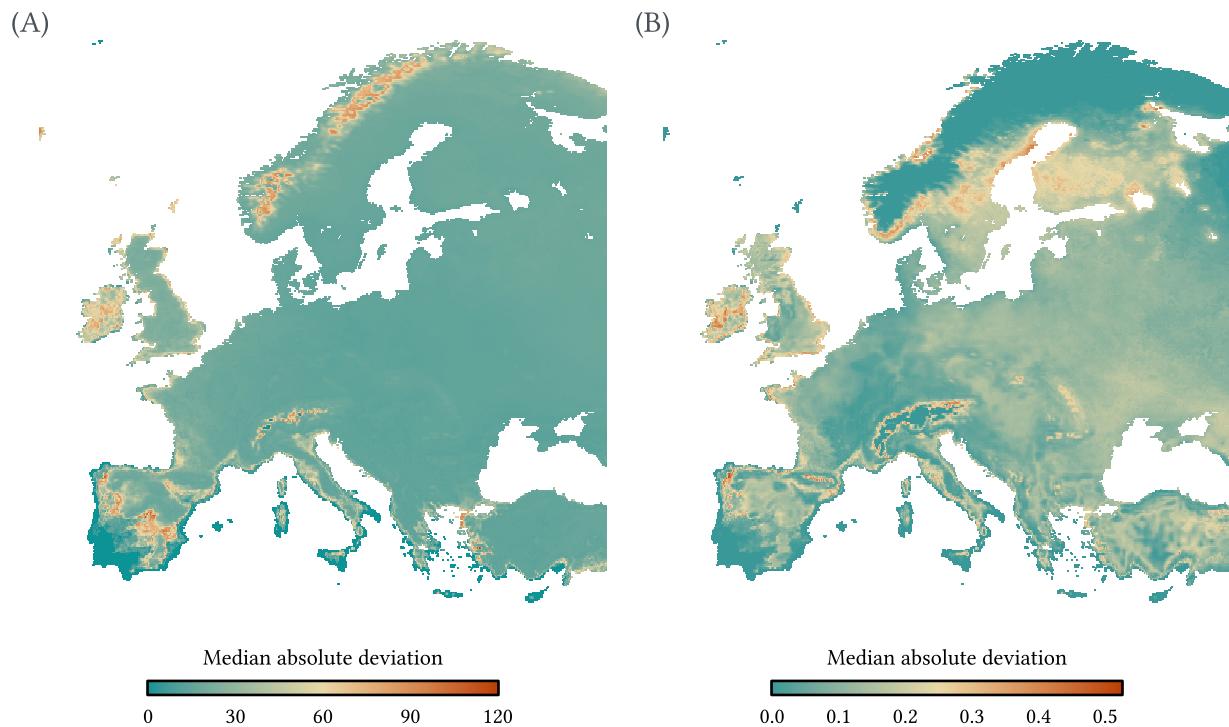


Figure G.2: Median absolute deviation of beech (A) leaf unfolding date and (B) fitness, predicted with 100 calibrated parameter sets of PHENOFIT.

679 The median standard deviation of unfolding date across Europe was about 15.4 days. On beech presence
 680 points, it was about 16.2 days. Nearly 90.3% of cells had a median absolute deviation lower than 30
 681 days ([Figure D.2.A.](#)). The median standard deviation of fitness across Europe was about 0.148. On
 682 beech presence points, it was about 0.153. Nearly 46.4% and 91.5% of total cells had a median absolute
 683 deviation lower than 0.1 and 0.2 respectively ([Figure D.2.B.](#)).

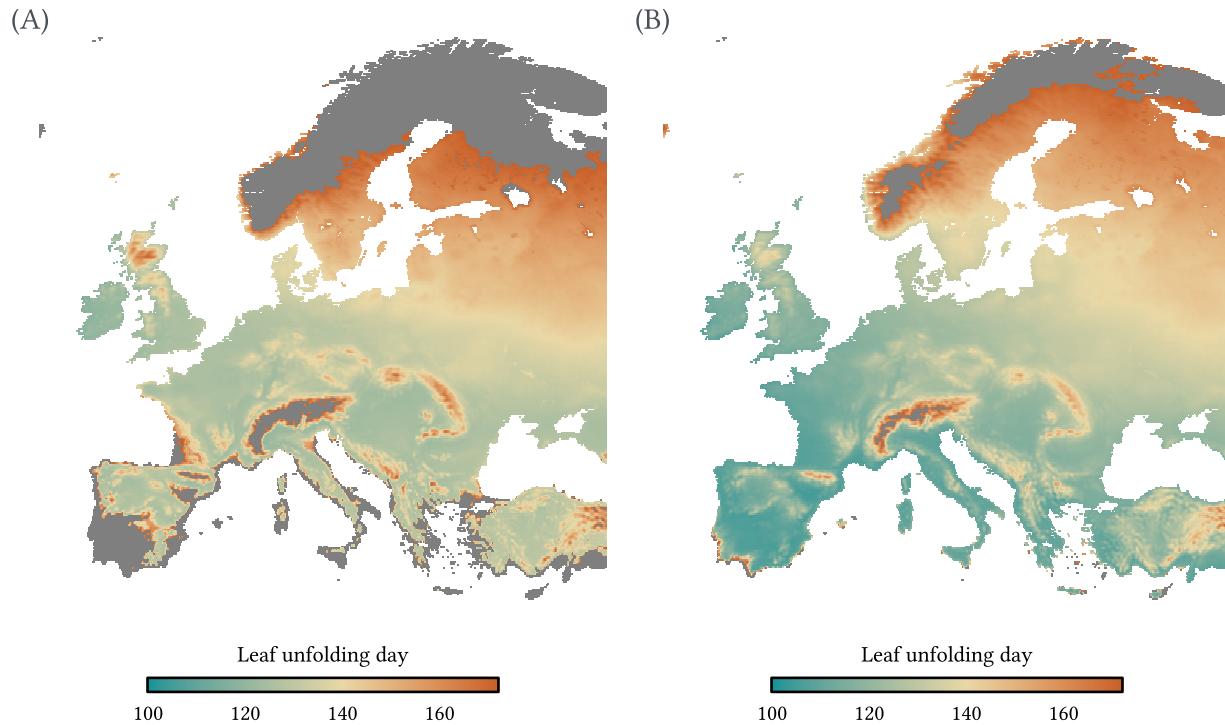


Figure G.3: Mean leaf unfolding day of beech with **(A)** best CMA-ES calibrated parameters and **(B)** classical (forward) parameters. Values above June solstice day are in grey. Note that PHENOFIT model assign a value of 365 when unfolding has not happened at all due to climate conditions.