

WILEY

ECOSPHERE

Tree regeneration in models of forest dynamics: a key priority for further research

Journal:	<i>Ecosphere</i>
Manuscript ID	ECS23-0342.R1
Wiley - Manuscript type:	Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Díaz-Yáñez, Olalla; ETH Zurich, Forest Ecology, Department of Environmental Systems Science Institute of Terrestrial Ecosystems Käber, Yannek; ETH Zurich, Forest Ecology, Department of Environmental Systems Science Institute of Terrestrial Ecosystems Anders, Tim; Senckenberg Biodiversity and Climate Research Centre Bohn, Friedrich; Environmental Research Center Leipzig, Department of Computational HydroSystem Braziunas, Kristin; Technical University of Munich, School of Life Sciences Brůna, Josef; Institute of Botany Czech Academy of Sciences Fischer, Rico; Helmholtz Center for Environmental Research UFZ, Department of Ecological Modelling Fischer, Samuel; Helmholtz-Centre for Environmental Research - UFZ, Department of Ecological Modelling Hetzer, Jessica; Senckenberg Biodiversity and Climate Research Centre Hickler, Thomas; Senckenberg Gesellschaft für Naturforschung, Hochauer, Christian ; University of Natural Resources and Life Sciences Vienna, Institute of Silviculture, Department of Forest and Soil Sciences Lexer, Manfred; University of Natural Resources and Life Sciences Vienna, Institute of Silviculture, Department of Forest and Soil Sciences Lischke, Heike; Swiss Federal Institute for Forest Snow and Landscape Research WSL, Dynamic Macroecology Mairotta, Paola; University of Bari Aldo Moro, Department of Soil Sciences, of Plants and Food Merganic, Jan; Technical University of Zvolen, Department of Forest Harvesting, Logistics and Ameliorations, Faculty of Forestry Merganičová, Katarina; Technicka Univerzita vo Zvolene Fakulta ekologie a environmentalistiky, Faculty of Forestry Mette, Tobias; LWF Bayern, Mina, Marco; EURAC Research, Institute for Alpine Environment Morin, Xavier; Centre National de la Recherche Scientifique, CEFE Nieberg, Mats; Potsdam-Institut für Klimafolgenforschung (PIK) e V Rammer, Werner; Technical University of Munich School of Life Sciences Weihenstephan Reyer, Christopher; Potsdam Institute for Climate Impact Research (PIK) e V Scheiter, Simon; Senckenberg Biodiversity and Climate Research Centre Scherrer, Daniel; Swiss Federal Institute for Forest Snow and Landscape Research, Bugmann, Harald; ETH Zurich, Forest Ecology, Department of

	Environmental Systems Science Institute of Terrestrial Ecosystems
Substantive Area:	Forestry < Ecosystems < Substantive Area, Modeling (general) < Statistics and Modeling < Theory < Substantive Area
Organism:	Plants
Habitat:	Deciduous Forest < Temperate Zone < Terrestrial < Habitat, Coniferous Forest < Temperate Zone < Terrestrial < Habitat, Mixed Forest < Temperate Zone < Terrestrial < Habitat
Geographic Area:	Europe < Geographic Area
Key words/phrases:	tree regeneration, tree recruitment, tree establishment, forest dynamics, models, landscape, stand, global, ingrowth, climate change
Abstract:	<p>Tree regeneration is a key process in forest dynamics, particularly in the context of forest resilience and climate change. Models are pivotal for assessing long-term forest dynamics, and they have been in use for more than 50 years. However, there is a need to evaluate their capacity to accurately represent tree regeneration. We assess how well current models capture the overall abundance, species composition, and mortality of tree regeneration.</p> <p>Using 15 models built to capture long-term forest dynamics at the stand, landscape, and global levels, we simulate tree regeneration at 200 sites representing large environmental gradients across Central Europe. The results are evaluated against extensive data from unmanaged forests.</p> <p>Most of the models overestimate recruitment levels, which is compensated only in some models by high simulated mortality rates in the early stages of individual tree dynamics. Simulated species diversity of recruitment generally matches observed ranges. Models simulating higher stand-level species diversity do not feature higher species diversity in the recruitment layer. The effect of light availability on recruitment levels is captured better than the effects of temperature and soil moisture, but patterns are not consistent across models. Increasing complexity in the tree regeneration modules is not related to higher accuracy of simulated tree recruitment. Furthermore, individual model design is more important than scale (stand, landscape, global) and approach (empirical, process-based) for accurately capturing tree regeneration.</p> <p>Despite the mismatches between simulation results and data, it is remarkable that most models capture the essential features of the highly complex process of tree regeneration, while not having been parameterized with such data. We conclude that much can be gained by evaluating and refining the modeling of tree regeneration processes. This has the potential to render long-term projections of forest dynamics under changing environmental conditions much more robust.</p>

1 *Journal: Ecosphere. Manuscript ID: ECS23-0342. Manuscript type: Article*

2 ***Tree regeneration in models of forest dynamics: a key priority for***

3 ***further research***

4 Olalla Díaz-Yáñez^{*1}, Yannek Käber¹, Tim Anders²,

5 Friedrich Bohn³, Kristin H. Braziunas⁴, Josef Brůna⁵,

6 Rico Fischer⁶, Samuel M. Fischer⁶, Jessica Hetzer²,

7 Thomas Hickler², Christian Hochauer⁷, Manfred J. Lexer⁷,

8 Heike Lischke⁸, Paola Mairotta⁹, Ján Merganič¹⁰,

9 Katarina Merganičová^{11,12}, Tobias Mette¹³, Marco Mina¹⁴,

10 Xavier Morin¹⁵ Mats Nieberg,¹⁶ Werner Rammer⁴,

11 Christopher P.O. Reyer¹⁶, Simon Scheiter², Daniel Scherrer¹⁷,

12 Harald Bugmann¹

13 ¹ETH Zurich. Department of Environmental Systems Science. Institute of Terrestrial Ecosystems.
14 Universitätstrasse 16, 8092 Zürich, Switzerland.

15 ²Senckenberg Biodiversity and Climate Research Centre (SBiK-F), 60325 Frankfurt am Main,
16 Germany.

17 ³Department of Computational Hydrosystem, Helmholtz-Centre for Environmental Research -
18 UFZ, Permoserstraße 15, 04318 Leipzig, Germany.

19 ⁴Ecosystem Dynamics and Forest Management Group, TUM School of Life Sciences, Technical
20 University of Munich, 85354 Freising, Germany.

21 ⁵Department of Geoecology, Institute of Botany of the Czech Academy of Sciences, Průhonice,
22 Czech Republic.

23 ⁶Department of Ecological Modelling, Helmholtz-Centre for Environmental Research - UFZ,
24 Permoserstraße 15, 04318 Leipzig, Germany.

25 ⁷Institute of Silviculture, Department of Forest and Soil Sciences, University of Natural Resources
26 and Life Sciences, Peter Jordan Strasse 82, A-1190 Vienna, Austria.

27 ⁸Dynamic Macroecology, Swiss Federal Institute for Forest, Snow and Landscape Research WSL,
28 Zürcherstrasse 111, 8903 Birmensdorf, Switzerland.

29 ⁹Department of Soil Sciences, of Plants and Food, University of Bari Aldo Moro, Italy.

30 ¹⁰Department of Forest Harvesting, Logistics and Ameliorations, Faculty of Forestry, Technical
31 University in Zvolen, Slovakia.

32 ¹¹Czech University of Life Sciences Prague, Faculty of Forestry and Wood Sciences, Kamýcká
33 129, 16500 Praha 6 – Suchdol, Czech Republic.

34 ¹²Department of Biodiversity of Ecosystems and Landscape, Slovak Academy of Sciences,
35 Štefánikova 3, P.O.Box 25, 81499 Bratislava, Slovakia.

³⁶ ¹³Dpt. of Soil and Climate, Bavarian State Institute of Forestry, Germany.

³⁷ ¹⁴Institute for Alpine Environment, Eurac Research, Bolzano/Bozen, Italy.

³⁸ ¹⁵CEFE, CNRS Campus du CNRS 1919, route de Mende 34293 Montpellier cedex, France.

³⁹ ¹⁶Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association,
⁴⁰ 14473 Potsdam, Germany.

⁴¹ ¹⁷Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111,
⁴² 8903 Birmensdorf, Switzerland.

⁴³*Corresponding author (olalla.diaz@usys.ethz.ch)

⁴⁴ **OPEN RESEARCH STATEMENT:** This submission uses novel code, both the data and code are
⁴⁵ provided in an external repository to be evaluated during the peer review process. The repository
⁴⁶ can be downloaded from [this figshare private link](https://figshare.com/s/100857ba67d2351dcb15): <https://figshare.com/s/100857ba67d2351dcb15>.

⁴⁷ The protocol is available at Díaz-Yáñez, Käber, and Bugmann (2022) ([with link](https://osf.io/czdxp/): <https://osf.io/czdxp/>).
⁴⁸ We will make publicly available the code and data if the manuscript is accepted.

⁴⁹ **Keywords:** tree regeneration | tree recruitment | tree establishment | forest dynamics | models |
⁵⁰ landscape | stand | global

51 Abstract

52 Tree regeneration is a key process in forest dynamics, particularly in the context of forest resilience
53 and climate change. Models are pivotal for assessing long-term forest dynamics, and they have been
54 in use for more than 50 years. However, there is a need to evaluate their capacity to accurately
55 represent tree regeneration. We assess how well current models capture the overall abundance,
56 species composition, and mortality of tree regeneration.

57 Using 15 models built to capture long-term forest dynamics at the stand, landscape, and global
58 levels, we simulate tree regeneration at 200 sites representing large environmental gradients across
59 Central Europe. The results are evaluated against extensive data from unmanaged forests.

60 Most of the models overestimate recruitment levels, which is compensated only in some models by
61 high simulated mortality rates in the early stages of individual tree dynamics. Simulated species
62 diversity of recruitment generally matches observed ranges. Models simulating higher stand-level
63 species diversity do not feature higher species diversity in the recruitment layer. The effect of light
64 availability on recruitment levels is captured better than the effects of temperature and soil mois-
65 ture, but patterns are not consistent across models. Increasing complexity in the tree regeneration
66 modules is not related to higher accuracy of simulated tree recruitment. Furthermore, individ-
67 ual model design is more important than scale (stand, landscape, global) and approach (empirical,
68 process-based) for accurately capturing tree regeneration.

69 Despite the mismatches between simulation results and data, it is remarkable that most models
70 capture the essential features of the highly complex process of tree regeneration, while not having
71 been parameterized with such data. We conclude that much can be gained by evaluating and refining

72 the modeling of tree regeneration processes. This has the potential to render long-term projections
73 of forest dynamics under changing environmental conditions much more robust.

For Review Only

74 Introduction

75 Forests provide a wide range of products and services of vital importance to humankind ([FOR-](#)
76 [EST EUROPE 2020](#)). Under the pressure of climate change, increasing disturbance impacts and
77 changing societal demands on forest ecosystem services, it is becoming ever more important to
78 understand how forest structure, composition and function will change, and to evaluate forest ca-
79 pacity to adapt to or be resilient after disturbance ([Lindner et al. 2010; Seidl and Turner 2022](#)). A
80 wide range of models of forest dynamics have been developed over the past decades considering
81 the impacts of climate ([Bugmann and Seidl 2022](#)). From these studies, it is evident that we have
82 a reasonably good understanding how to model tree growth ([Vanclay and Skovsgaard 1997; Bug-](#)
83 [mann et al. 1996](#)), and substantial efforts have been dedicated to improving the representation of
84 tree mortality ([Bugmann et al. 2019; Cailleret et al. 2017](#)). In contrast, tree regeneration is much
85 less studied and is often represented rather coarsely in models ([Walck et al. 2011; Leishman et al.](#)
86 [1992; Price et al. 2001](#)). This presents an important research gap, particularly in the context of
87 climate-induced forest disturbances and forest resilience.

88 Tree regeneration arises from multiple processes including pollination, fruit maturation, seed pro-
89 duction, dispersal, germination, juvenile growth and survival ([Price et al. 2001; Vacchiano et al.](#)
90 [2018](#)). All these processes are difficult to assess, and some of them are scarcely understood and thus
91 appear highly stochastic ([Bogdziewicz et al. 2021](#)). Identifying the appropriate level of complexity
92 for the mathematical formulation of the key factors that are leading to successful tree regeneration
93 is challenging. Currently, tree regeneration processes in dynamic forest models are handled in a
94 multitude of ways ([König et al. 2022; Bugmann and Seidl 2022](#)): from 1) entirely ignoring them
95 (as done in classical forest growth models, e.g., Pretzsch et al. [2002](#)), across 2) the use of a few

96 simple environmental filters, as done in most forest “gap” models ([Shugart 1984](#)) and Dynamic
97 Global Vegetation Models (e.g., [Hickler et al. 2012](#); [Smith, Prentice, and Sykes 2001a](#)), to 3) com-
98 plex approaches that incorporate local feedback from the canopy, multiple ecological processes
99 and often also short time steps (e.g., [Seidl et al. 2012](#); [Wehrli et al. 2006](#)), or 4) field-based statisti-
100 cal parameterizations, which however are not easy to extrapolate in space and time (e.g., [Ribbens,](#)
101 [Silander, and Pacala 1994](#)).

102 Overall, models are needed to (1) synthesize existing empirical data and explore their relationships,
103 (2) assess future tree regeneration e.g. in the context of global change scenarios, and (3) identify
104 the most important processes that are shaping vegetation patterns. Given the strategies that are
105 used in models of forest dynamics to represent tree regeneration, their behavior often is prone to
106 problems, such as very high levels of tree regeneration that necessitate excess mortality at early
107 stages of tree life to simulate correct stand structure and composition ([Kroiss and HilleRisLambers](#)
108 [2015](#)). Also, inadequately high species diversity in tree regeneration may be simulated, which
109 is characteristic of “classical” forest gap models ([Gutiérrez, Snell, and Bugmann 2016](#)), at least as
110 long as the simulation set-up comprises a multitude of species. Some models use calibration against
111 local canopy-level data to constrain simulated tree regeneration, which is likely to hamper the
112 general applicability of these models e.g. under scenarios of climate change. Furthermore, correctly
113 capturing the species composition of tree regeneration as a function of the presence of seed trees in
114 the canopy is often a particular challenge, potentially leading to unrealistic successional drift in the
115 model, which must be corrected by factors that are hard to parameterize (e.g., [Lischke and Löffler](#)
116 [2006](#)). A related issue is the excessive reduction of species diversity due to positive feedback
117 effects, such that eventually just single-species stands remain ([Meier et al. 2011](#)). This is sometimes

118 corrected by the incorporation of a low level of seed influx of all species at all times (Schumacher
119 et al. 2006) or by restricting the number of seeds per species in the seedbank (Lischke and Löffler
120 2006). However, simulated species composition is usually exceedingly sensitive to assumptions
121 about seed availability, whereas the parameters of such functions are poorly constrained by field
122 data. Lastly, there is often a problem with insufficient observational constraints on parameter values
123 for models that start from very small tree sizes (e.g., 10 cm in height, or even from seed) and track
124 tree development in a process-oriented manner by considering a multitude of ecological influences,
125 rather than emphasizing tree recruitment into a larger size class (e.g., Wehrli et al. 2006).

126 Modeling tree regeneration processes is challenging, and even in empirical ecology, it has not re-
127 ceived much attention (Hanbury-Brown, Ward, and Kueppers 2022), although many sub-processes
128 have been studied in detail (Miina, Eerikäinen, and Hasenauer 2006). Yet, little data are avail-
129 able that cover all the processes within one species along environmental gradients, let alone for a
130 vast suite of species. Data on forest regeneration are often fragmented, which constitutes a ma-
131 jor problem for model building (Clark et al. 1999). For example, monitoring on permanent plots
132 (such as National Forest Inventories) often measures the ingrowth of new trees into a specific size
133 class (Zell et al. 2019), however with a design that captures the rate after the stand initiation phase
134 (Hallsby et al. 2015). Data from permanent plots, although available across regions, are also highly
135 heterogeneous, rendering their use difficult in a modeling context (Käber et al. 2021). Lastly, tar-
136 geted studies to measure tree regeneration on experimental sites are often limited in spatial extent,
137 species studied, or the sub-set of processes that are investigated (Berdanier and Clark 2016; Collet
138 and Chenost 2006).

139 Thus, a focus on the modeling of tree regeneration processes is sorely needed and overdue (Price

140 et al. 2001; Walck et al. 2011) if we are to make reliable projections of future forest dynamics, i.e.,
141 when the models need to be operated in extrapolation mode, as well as from a fundamental ecolog-
142 ical point of view for increasing systems understanding. In the present study, for the first time, a
143 large number of forest models commonly used to assess forest dynamics under climate change are
144 evaluated against a continental-scale, multi-species harmonized dataset on tree recruitment (Käber
145 et al. 2023). By tree recruitment, we refer to the passing of trees across a specific diameter thresh-
146 old (“ingrowth”). We included models that are based on a range of “philosophies”, from models
147 operating at the stand to the global scale as well as the range of models from empirically derived
148 to “process-based” (e.g., Bugmann et al. 1996; Fabrika and Ďurský 2012; Lexer and Hönniger
149 2001; Reyer et al. 2014). Due to the large variability in tree regeneration patterns in nature and
150 the large number of factors driving and constraining this process — including some that are not
151 incorporated explicitly in most models, such as deer browsing — we do not aim for a detailed sta-
152 tistical evaluation of each model. Instead, we aim to evaluate the general recruitment patterns and
153 magnitudes simulated by the models, and benchmark the simulated regeneration niche of multiple
154 species against empirical data along a wide environmental gradient of temperature, moisture, and
155 light availability (Grubb 1977).

156 More specifically, we aim to answer the following questions: (1) Are models of forest dynamics
157 capturing accurately tree recruitment levels, initial tree species diversity and mortality in the recruit-
158 ment? (2) Do model traits explain differences in model performance? (3) How well do the models
159 capture total recruitment and the regeneration niches of individual species across environmental
160 gradients of light availability, temperature, and soil moisture?

161 We evaluate the performance of the models by analyzing simulated data alongside observed data.

¹⁶² Specifically, we examine tree species diversity, mortality rates in tree recruitment, and recruitment
¹⁶³ levels. Then, we link model traits (e.g., complexity, scale) to performance, and we assess whether
¹⁶⁴ modeled total recruitment or individual species regeneration niches (i.e., proportional recruitment
¹⁶⁵ along an environmental gradient) align with observational data.

¹⁶⁶ Material and methods

¹⁶⁷ Models

¹⁶⁸ Fifteen models of forest dynamics (including two models featuring model variants) developed for
¹⁶⁹ the stand, landscape or global scale were used to simulate forest dynamics (Table 1). The ap-
¹⁷⁰ proaches used for model construction and their origin differ strongly, with most of the models fea-
¹⁷¹ turing a largely “process-based” approach, whereas two models are based on formulations derived
¹⁷² from the statistical analysis of inventory data (SIBYLA and xComp). Some of the models largely
¹⁷³ rely on the original approach underlying forest gap models (e.g., ForClim 1), whereas some are
¹⁷⁴ based on plant ecophysiological processes (e.g., FORMIND and iLand). The different approaches
¹⁷⁵ underlying the models have strongly influenced the formulation of tree regeneration processes.

¹⁷⁶ The models can be differentiated into ‘regeneration’ and ‘recruitment’ approaches ([Vanclay and](#)
¹⁷⁷ [Jerome 1994; König et al. 2022](#)). *Regeneration models* include processes such as flowering and
¹⁷⁸ pollination, seed production, seed dispersal, germination, and seedling growth, which ultimately
¹⁷⁹ lead to the simulated number of established trees. *Recruitment models*, in contrast, introduce a num-
¹⁸⁰ ber of new trees with certain characteristics such as biomass or diameter, without explicitly consid-
¹⁸¹ ering earlier development processes. We can further distinguish models that feature a feedback from

¹⁸² stand properties to simulated recruitment, i.e., where the level and species composition of recruitment
¹⁸³ is influenced by the adult tree community via the production of seeds (or seedlings/saplings),
¹⁸⁴ from models that do not contain such feedback.

¹⁸⁵ The starting point for tree regeneration in the models differs as well, ranging from seedbank, seed
¹⁸⁶ or seedling to sapling (i.e., trees much larger than 10 cm height, often ca. 2-3 m tall). Models
¹⁸⁷ that start from seed need to include a larger number of ecological processes such as germination
¹⁸⁸ and survival of young seedlings, whereas models that start from saplings have to aggregate via
¹⁸⁹ parameterizations several ecological processes that are not treated explicitly. This latter approach
¹⁹⁰ reduces model complexity but comes at the cost of blurred process representation.

¹⁹¹ In the models used here, the overall complexity in the regeneration modules varies considerably.
¹⁹² Following Bugmann and Seidl (2022), we can classify seven models (iLand, PICUS, LandClim,
¹⁹³ ForCEEPS, LPJ-GUESS, ForClim and TreeMig) as having rather high complexity in their re-
¹⁹⁴ generation modules (mean regeneration formulation complexity across all processes >0, Table
¹⁹⁵ 1), whereas the other eight models feature relatively low complexity. Two models, ForClim and
¹⁹⁶ ForCEEPS, were used here with two alternative variants of regeneration. ForClim variant 1 (Bug-
¹⁹⁷ mann et al. 1996) is based on a regeneration module that adheres closely to the concept introduced
¹⁹⁸ by Botkin, Janak, and Wallis (1972) which is based on environmental filters considering species-
¹⁹⁹ specific thresholds of light availability and climatic variables. In contrast, ForClim variant 11 is
²⁰⁰ adopting a slightly more complex approach where individual species properties and their relation
²⁰¹ to the environment are represented gradually along with the relative suitability of different species
²⁰² (i.e., competition among tree species) (Huber, Bugmann, and Lafond 2020). These two model vari-
²⁰³ ants allow us to evaluate a more process-based and complex module (variant 11) against a simple

204 module (variant 1), while the rest of the model structure is identical. Similarly, the two ForCEEPS
205 variants allow us to isolate the importance of the canopy feedback (i.e., simulated actual composi-
206 tion and relative abundance of species in the plot) via seed trees for the quantity and quality (e.g.,
207 diversity and composition) of simulated regeneration, as one variant includes this feedback whereas
208 the other does not.

209 Observed data

210 Recruitment data covering a wide range of environmental conditions are hard to obtain, and this
211 is one of the reasons why most models of forest dynamics have never been confronted with a
212 dataset covering such gradients over a large number of sites, to evaluate how well regeneration is
213 captured. The observations used here are derived from a novel network of sites in forest reserves
214 that represent the range of environmental gradients in temperature and precipitation in Central
215 Europe as compiled in the EuFoRIa network (EuFoRIa 2019; Käber et al. 2023) (Figure 1). The
216 environmental coverage and the variety of forests represented by this data is unprecedented within
217 Europe. These forests have been unmanaged for at least 10 years prior to their designation as
218 reserves, and most of them provide time series of natural forest dynamics over multiple decades
219 and up to ca. 85 years. The census periods range from 3 to 37 years, with an average of 14
220 years. The data provide information at the tree level, thus allowing for the sequential comparison
221 of processes such as individual-tree recruitment and death between the measurements.

222 We selected 200 sites from this network as the benchmarking dataset for the simulation, so as
223 to be representative of the environmental variation contained in the data. This was achieved by
224 applying k-means clustering to define 200 clusters of plots from the original set of 869 plots along

225 the environmental dimensions of temperature, climatic water balance, soil quality, slope and aspect.
226 Each of these sites featured at least two consecutive measurements. Recruitment size thresholds
227 ranged from stem diameters at breast height (dbh) from 0 to 10 cm. For our study, we defined two
228 datasets, where one included 165 sites with a diameter threshold of 7 cm or lower, and the other
229 included another 35 sites with diameter thresholds between 7 and 10 cm. Therefore we had in total
230 200 sites with observations of newly recruited trees above 10 cm and a subset of 165 sites with
231 recruitment data above 7 cm. In the results section, any observations from the data at the 7 cm
232 threshold originated from 165 sites.

233 Plot size ranged from 0.02 ha to 5.52 ha, and the data were further processed and aggregated fol-
234 lowing Käber et al. (2023) to provide recruitment rates per ha and per decade. The observations
235 used in this study featured 30,900 newly established trees. Recruitment rates per site, sample and
236 decade ranged from zero to 1,246 trees, with a mean of 56 trees. Adult species composition was also
237 available for each plot. For more details on this dataset and the detailed recruitment information,
238 cf. Käber et al. (2023).

239 **Simulation protocol**

240 The overarching goal of the simulation experiments was to assess tree recruitment as it arises from
241 empirical data against its representation in a wide range of models of forest dynamics. We define
242 tree recruitment as the passing of a breast height diameter threshold of 7 and 10 cm (synonym:
243 ingrowth). To this end, each modeling group was provided with a detailed protocol (Díaz-Yáñez,
244 Käber, and Bugmann 2022) with instructions how to perform the simulations, input variables on
245 climate and soil conditions, and the list of expected output variables. Neither were further site

246 information (except for the data specified below) nor any data on tree recruitment or forest stand
247 features provided prior to the simulation.

248 The input variables were collected from different data sources and aggregated to be suitable for the
249 needs of the different models. Time series of climatic variables from 1981 to 2018 were provided in
250 hourly (Era5-land data, Muñoz Sabater 2019), daily and monthly resolution (CHELSA data, Karger
251 et al. 2021). Some variables required for some models, such as relative humidity or vapor pressure
252 deficit, were calculated from these variables. The final instructions for using the climate data were
253 slightly different in each model, based on the approach that best suited the model (Table 1). Soil
254 quality data were provided as continuous values between 1 and 5 (Soilgrids dataset, Hengl et al.
255 2017). The protocol also provided the elevation, slope and aspect for each of the 200 sites (ASTER
256 Science Team 2019), but no other spatial information such as coordinates, with the exception of
257 iLand and aDGVM2, which required blurred coordinates to derive highly detailed soil data.

258 The simulations were run in the absence of natural disturbances. Only the model LPJ-GUESS had
259 to include a “background” disturbance probability to increase chances of shade-intolerant species
260 to establish. The simulations were set up to sample species-specific recruitment rates per decade
261 and per ha in the equilibrium state of the model, typically entailing a “spin-up” run (as we did
262 not provide any forest data). The modeling teams decided on the simulated area and how they
263 derived these samples. The exact length of the simulation was also decided by the modeling teams
264 (Table 1). Further details on how each modeling team prepared the simulations and the outputs are
265 available in Appendix S1: Model reports.

266 The simulations were run in the absence of management to a simulated equilibrium (“Potential
267 Natural Vegetation”) with the current climate, as the ultimate goal was to evaluate tree recruit-

ment under comparable and near-equilibrium conditions. This entails the assumption that (1) the observations from the forest reserves reflect no traces of forest management, and (2) there is an equilibrium between forest dynamics and climate. While the former might be starting to be visible in many of the EuFoRIA reserves, the latter may be more debatable. However, in the absence of detailed data on the history of each plot in the EuFoRIA network, some broad assumptions had to be made. Both the width of the regeneration niche (i.e., in environmental space) as well as the intensity of recruitment (i.e., the number of recruited trees per area and per unit of time) were of interest.

The simulations were run for mixed-species forests (not multiple single-species simulations) using mixtures of eleven species or genera for which regeneration data of sufficient quality were available from EuFoRIA: *Fagus sylvatica* L., *Picea abies* L., *Abies alba* Mill., *Carpinus betulus* L., *Tilia cordata* Mill., *Acer pseudoplatanus* L., *Betula* spp. L., *Fraxinus excelsior* L., *Quercus* spp. L., *Alnus glutinosa* L., and *Pinus sylvestris* L. The same set of species was used at all 200 sites. Two models included their standard set of species for the simulations, which is much larger (i.e., ForClim 1, ForClim 11, TreeMig), and the species simulated beyond the eleven species were lumped in a category “others”. In three models, fewer than these eleven species were simulated (4C, xComp, LPJ-GUESS) (Table 1). In 4C only *Fagus sylvatica*, *Picea abies*, *Betula* spp., *Quercus* spp. and *Pinus sylvestris* are parameterised. xComp simulations did not consider *Fraxinus excelsior* due to a deprecated species parameterisation. In LPJ-GUESS, *Acer pseudoplatanus* and *Alnus glutinosa* are not parameterized and therefore these species could not be included in the simulation. Finally, the model aDGVM2 does not simulate individual species. Rather, community assembly processes and trait filtering generate plant communities that are adjusted to the local biotic and abiotic conditions,

290 and the simulated plants can be classified into ecological strategies based on their trait values in a
291 post-processing step.

292 Each of the models reported the number of new trees crossing the two diameter thresholds by
293 sampling 200 times in a 10-year interval for each species and per ha for each of the 200 sites.

294 Multiple samples per site were used to better understand the simulated variation within each site.

295 This was done using different strategies, depending on the model, including (1) sampling simulated
296 data from the same 1 ha plot in the equilibrium over time, (2) sampling several 1 ha plots from the
297 simulated forest at one specific point in time (in the equilibrium) or (3) a combination of (1) and

298 (2). This resulted in 880,000 observations per model that simulated the 11 species included in the
299 protocol (200 sites, 200 samples per site, 11 species and two diameter thresholds). For the models
300 that simulated additional species, their regeneration rates were aggregated as “others”.

301 Two models did not provide results from all the simulated samples or sites to avoid unrealistic re-
302 sults, as follows. In the model 4C, a threshold of a maximum basal area of $90\ m^2\ ha^{-1}$ was used

303 to avoid unrealistic stand basal area data, and therefore not all the samples and sites were reported.

304 The reason for this is that 4C is not suitable for long-term simulations without management, due
305 to misrepresentations in density-dependent mortality processes in long-term simulations and as-
306 sumptions of tree geometry that lead to unrealistic single tree dimensions of very old individuals.

307 LPJ-GUESS had 2% of the sample outputs without tree recruitment, and these were considered as
308 zero stand basal area and zero regeneration for all the tree species simulated; one site produced
309 grassland rather than a forest and was not included in the results.

³¹⁰ **Data analysis**

³¹¹ The simulation results were analyzed regarding (1) recruitment levels (i.e., ingrowth number per
³¹² unit time and space), (2) recruitment species diversity, (3) recruitment mortality, (4) the relationship
³¹³ between model performance and model traits, and (5) ingrowth gradients along the regeneration
³¹⁴ niches. We evaluated species diversity across the models and in relation to the observed data by
³¹⁵ calculating the Shannon index H based on the relative proportion of the species in terms of basal
³¹⁶ area. It was calculated for the regeneration (H_{R_n}) (Equation 1) and at the stand level ($H_{S_{n,s}}$)
³¹⁷ (Equation 3). The higher the value of the index, the higher the species diversity at a particular site
³¹⁸ and sample.

$$H_{R_n} = - \sum_{i=1}^s p_{R_{n,i}} \ln(p_{R_{n,i}}) \quad (1)$$

$$p_{R_{n,i}} = \frac{r_{BA_{n,i}}}{Total_{BA_n}} \quad (2)$$

$$H_{S_n} = - \sum_{i=1}^s p_{S_{n,i}} \ln(p_{S_{n,i}}) \quad (3)$$

$$p_{S_{n,i}} = \frac{BA_{n,i}}{Total_{BA_n}} \quad (4)$$

³¹⁹ where s is the total number of species present that have a basal area larger than zero in sample n ;
³²⁰ $p_{R_{n,i}}$ is the proportion of species i in sample n calculated as the regeneration basal area ($r_{BA_{n,i}}$)

321 for that species i relative to total recruited basal area ($TotalR_{BA_n}$) of the sample n ; $p_{S_{n,i}}$ is the
322 proportion of species i calculated as the basal area of all trees ($BA_{n,i}$) of that species i relative to
323 total basal area ($Total_{BA_n}$) of the sample n . Species diversity was not assessed for aDGVM2 as
324 this model does not simulate individual species.

325 Mortality in tree recruitment was assessed based on the ratio of recruitment between the 7 and 10
326 cm diameter thresholds. We used the Reineke self-thinning rule (Reineke 1933; Hans Pretzsch
327 and Biber 2003) as a reference to estimate whether the ratio of recruitment between these two
328 thresholds was above or below the expected theoretical rate. The Reineke self-thinning rule is
329 usually calculated for even-aged, single species stands and is based on a fixed relationship between
330 the number of stems and the quadratic mean diameter in fully stocked pure stands. The value used
331 in our comparisons was 1.77 (i.e., we expect stem numbers at 7 cm to be 77% higher than at 10
332 cm), calculated using Equation 5.

$$Reineke = \frac{N7}{N10}^{-1.605} \quad (5)$$

333 We assessed model performance in relation to model traits focusing on (1) model complexity as
334 defined by Bugmann and Seidl (2022), (2) model type (empirical or process-based), (3) the pres-
335 ence or absence of a canopy feedback for regeneration, and (4) the scale of application of the model
336 (stand, landscape or global). We tested for significant differences between simulated and observed
337 values of the recruitment levels and species diversity by using t-tests with the Bonferroni correc-
338 tion for multiple testing. Note that the t-tests does not present conclusive evidence for a specific
339 hypotheses, but rather facilitates managing the extensive number of comparisons discussed in our
340 study.

341 Total recruitment and the regeneration niches of the individual species were evaluated across the
 342 gradients of light availability, temperature, and soil moisture as captured in the data. Stand basal
 343 area (the basal area of all the trees in each sample above the recruitment threshold) was used as
 344 a proxy for light availability at the forest floor, the annual degree-day sum (Allen 1976; Fischlin,
 345 Bugmann, and Gyalistras 1995) as a proxy for growing season warmth, and the climatic water bal-
 346 ance as a proxy for soil moisture (Speich 2019). Recruitment values of each species were calculated
 347 as the mean across the 200 samples per site. The observed data were modeled using a Generalized
 348 Additive Model (Wood 2011) with a negative binomial distribution and restricted maximum like-
 349 lihood to better understand the relationship between the environmental gradients and the levels of
 350 observed regeneration, relative to the simulation results per model.

351 In order to analyze the regeneration niches across the climatic gradients, we focused on five com-
 352 mon tree species or genera: *Abies alba*, *Fagus sylvatica*, *Picea abies*, *Pinus sylvestris*, and *Quercus*
 353 spp. For these species, we calculated the share in the recruited basal area per site, where each site
 354 represents a part of the environmental space, as the mean across the available samples per site:

$$\bar{R}_{BA_{share_{t,i}}} = \frac{\bar{r}_{BA_{t,i}}}{\sum_{i=1}^s \bar{r}_{BA_{t,i}}} \quad (6)$$

355 where s is the total number of species simulated, and $\bar{r}_{t,i}$ is the mean basal area in the recruitment
 356 (subsequently referred to as “recruited basal area”) of species i at site t across the available samples
 357 at that site. Furthermore, regarding the share in the recruited basal area per species ($\bar{R}_{BA_{share_i}}$),
 358 we categorized this as zero when the recruited basal area of that species ($\bar{r}_{BA_{t,i}}$) was zero, but also
 359 when both the total recruited basal area ($\sum_{i=1}^s \bar{r}_{BA_{t,i}}$) and the recruited basal area of that species
 360 ($\bar{r}_{BA_{t,i}}$) equaled zero.

361 Results

362 Recruitment levels

363 Simulated levels of recruitment varied strongly across the 15 models and typically did not match the
364 levels found in the forest reserve data (Figure 2). Recruitment was overestimated in most models
365 for both the 7 and 10 cm diameter thresholds, with the exception of the empirical stand model
366 SIBYLA, the landscape model Landis-II and the global model aDGVM2, which estimated recruited
367 levels at the lower end of the plausibility interval of the observed data (Figure 2). The models with
368 the largest overestimation were the stand model PICUS and the landscape model TreeMig. For
369 most models, the variability of simulated recruitment levels across the 200 sites (visible from the
370 interquartile range in the box plots of Figure 2) was similar to or smaller than observed, with the
371 notable exception of PICUS, where simulated recruitment variability was much larger.

372 Both the observations and the simulated data had no recruitment in some samples and at some
373 sites (for details, cf. Table S1). The observed data had 4% of the samples with no recruitment.
374 Only three models had a larger proportion of no recruitment (4C, Landis-II and aDVM2). Two
375 models always simulated recruitment for both the 7 and 10 cm threshold (xComp and TreeMig),
376 i.e., they did not feature any zero values. The other ten models had a very low percentage of samples
377 with no recruitment (0.01% - 2.39%), i.e., they had distinctly fewer occurrences of zeros than the
378 observations.

379 Tree species diversity of recruitment

380 Most models matched the level of diversity of the observed data quite well (Figure 3). Five models
381 overestimated recruitment diversity: ForCEEPS, ForClim 1, PICUS, TreeMig and LandClim, the
382 latter particularly for the 7 cm diameter threshold. The model 4C is a special case, as it simulated
383 five species only, i.e., its diversity values are not directly comparable to those of the other models,
384 nor to the observations. Only one model, Landis-II, consistently underestimated recruited diversity.

385 In most models, there were only small but significant differences in the species diversity of recruit-
386 ment between 7 and 10 cm across sites (Appendix S1: Table S2). Four models (iLand, Landis-II,
387 TreeMig and LPJ-GUESS) maintained the recruitment diversity between the 7 and 10 cm thresh-
388 olds (i.e., the differences between them were not significant, see Appendix S1: Table S2), and the
389 same was evident from the observed data.

390 For both the observed and the simulated data, and both recruitment thresholds, species diversity
391 in the recruitment compared to stand-level diversity did not feature clear patterns across the mod-
392 els (Appendix S1: Figures S1 and S2). Most models captured reasonably well or overpredicted
393 species diversity at the stand level (Figure 4A, left and center; cf. Appendix S1: Figure S3), but a
394 characteristic feature was that the simulations had a much lower variability of diversity compared
395 to observations. Models overpredicting stand-level species diversity included ForClim1, ForClim
396 11, FORMIND and TreeMig (Figure 4A, left). Several models underpredicted species diversity,
397 i.e., LPJ-GUESS, ForCEEPS(f) and 4C (Figure 4A, right).

398 The majority of the models overpredicted recruitment diversity; as already observed in Figure 3,
399 only two of them underestimated it (4C and Landis-II) (Figure 4B and Appendix S1: Figure S4).
400 Both the models with and without feedback from the adult trees to regeneration (via seed pro-

401 duction) overestimated, underestimated or captured reasonably well recruitment species diversity
402 compared to the observations. Again, simulated diversity had considerably lower variability than
403 observed diversity in the recruitment.

404 Recruitment mortality

405 There were strong differences among models regarding the mortality rate between 7 and 10 cm
406 (Figure 5). Most of the models featured a mortality rate significantly larger than the observed
407 data, thus at least partially compensating for the general overestimation of recruitment levels at
408 7 cm (Figure 2 and Appendix S1: Table S3); six models did not feature a significant difference
409 compared to the observed data (ForCEEPS, FORMIND, ForClim 1 and 11, iLand and aDGVM2).

410 Four models (xComp, PICUS, LandClim and TreeMig) featured very high mortality rates (i.e., well
411 above the Reineke self-thinning line), which compensated for the strong initial overestimation of
412 recruitment (cf. Figure 2). Interestingly, two models that underestimated overall recruitment levels
413 (SIBYLA and Landis-II; cf. Figure 2) featured mortality rates that were close to but still above the
414 self-thinning line (Figure 5).

415 The fact that some models (and observed data) featured lower mortality than expected by Reineke's
416 self-thinning rule may point to facilitation, or simply a higher mortality rate before the trees had
417 reached 7 cm diameter. Yet, the case of models such as xComp, PICUS and TreeMig, whose mor-
418 tality was well above the self-thinning line, indicates that they feature higher recruitment mortality
419 compared to even-aged, single-species forests. Some models yielded a mortality rate of recruitment
420 that is broadly compatible with the self-thinning rule (e.g., 4C and LandClim).

421 The three models with the largest overestimation of the proportion of recruitment at the 7 cm thresh-

old featured the highest mortality between the 7 and 10 cm thresholds (Figure 6A). The other models that overestimated recruitment had mortalities lower than the expected self-thinning ratio, with the exception of LandClim (Figure 6B). It is noteworthy that most of the models that featured a low ratio of recruitment between 7 and 10 cm (i.e., ForCEEPS, ForClim 1, ForClim 11 and iLand) had only a small overestimation of recruitment at the 7 cm threshold (Figure 5). aDGVM2 was the only model that underestimated recruitment at the 7 cm threshold while featuring a low mortality between the two thresholds (Figure 6C), and only a few models were close to the correct recruitment levels at 7 cm while also being close to the Reineke line (SIBYLA, Landis-II; cf. Figure 6D; and 4C as well as LandClim).

Lastly, some models that had a low number of recruits at the 7 cm threshold also had a small decrease of recruitment between 7 and 10 cm (ratio 7-10 cm) (Appendix S1: Figure S5). In these models, competition and self-thinning are either not pronounced or must have occurred before the trees had reached 7 cm. However, this pattern was not consistent across models (Appendix S1: Figure S5). It is noteworthy that there is a relationship between the recruitment levels at 7 cm and the mortality rate (Appendix S1: Table S4): most of the models showed a positive effect (higher regeneration at 7 cm is coupled to an increase of the mortality between 7 and 10 cm). While some models showed a negative effect, the linear trend was not pronounced and not always significant regarding its slope.

Model performance and model traits

There was no significant relationship between the mean complexity of the regeneration module (Table 1) and the overestimation proportion at a diameter threshold of 7 cm (Appendix S1: Figure

443 S6). All models except 4C had significant differences between the observed and simulated mean
444 recruitment values. Therefore it was not possible to differentiate in terms of performance between
445 models type (empirical or process based), or their scales (stand, landscape or global) (Appendix
446 S1: Table S5). The values of species diversity in the recruitment simulated by each model were
447 also assessed against the observed data to see if there was a difference between models that include
448 a feedback compared to those that did not; all the models except ForCEEPS(f) and FORMIND had
449 significant differences (Appendix S1: Table S6). It is noteworthy that for the model ForCEEPS the
450 pattern changed from the overprediction of species diversity in the regeneration to a diversity level
451 that is closer to observations when the recruitment module included feedback (ForCEEPS(f)).

452 Regeneration gradients and regeneration niches

453 When evaluating total tree recruitment levels along key gradients of light availability (basal area),
454 temperature (degree-days) and soil moisture (climatic water balance), distinct features emerged:
455 the models reproduced the effect of basal area in both its magnitude and patterns (Figure 7 and
456 Appendix S1: Figure S7) considerably better than the effects of the climatic gradients, where they
457 featured varying patterns.

458 In the observed data, recruitment levels decreased clearly with increasing stand basal area (Figure
459 7a). This trend was captured by the Landis-II model only, albeit at much lower values of basal
460 area than in the observations. The other models featured distinctly different trends, such as (1) an
461 increase in recruitment levels with increasing total basal area followed by the absence of recruitment
462 at high values of basal area (ForCEEPS, ForCEEPS(f), FORMIND, TreeMig and LPJ-GUESS), (2)
463 almost constant recruitment levels with basal area (ForClim 11, xComp, LandClim and aDGVM2),

464 or (3) an increase of recruitment up to a certain value of stand basal area followed by a decrease
465 at even higher values of stand basal area, with model-specific thresholds (4C, ForClim 1, PICUS,
466 SIBYLA and iLand).

467 The observed recruitment did not change much across the climatic water balance gradient, and
468 showed a slight increase with temperature. Neither did the models generally match the observed re-
469 cruitment patterns across these gradients (Figure 7b,c), nor was there a clear pattern across models.

470 Regarding the soil moisture gradient, a group of models featured decreasing recruitment with in-
471 creasing soil moisture (4C, ForClim 1, ForClim11, SIBYLA, xComp, PICUS, LandClim, TreeMig
472 and iLand), while a few models showed the opposite trend (ForCEEPS(f), FORMIND and LPJ-
473 GUESS). Across the temperature gradient, all the models that featured a decrease in recruitment
474 levels with increasing soil moisture showed the opposite trend (i.e., more recruitment with increas-
475 ing temperature), except for 4C. The other models did not feature a clear pattern.

476 Silver fir (*Abies alba*), beech (*Fagus sylvatica*), Scots pine (*Pinus sylvestris*) and oak (*Quercus*
477 spp.) were the species observed most frequently in the EuFoRIa data. Most models captured well
478 the share of basal area in the recruitment of these five species compared to the observed data along
479 the two major climatic gradients (Figure 8, cf. Appendix S1: Figure S8). However, the models
480 differed considerably in the way in which they simulated these climatic niches.

481 First, some models overestimated recruitment across the entire gradients of at least one of the five
482 main tree species (Figure 8, e.g., ForCEEPS, ForCEEPS(f), LandClim or LPJ-GUESS), while other
483 models overestimated recruitment of more than one species for a large part of the climate gradients,
484 such as xComp, PICUS, SIBYLA or LPJ-GUESS (cf. red colors in Figure 8). These trends were
485 consistent for both recruitment thresholds, although the exact changes in the share of the recruited

486 basal area were different (Appendix S1: Figure S8).

487 Second, there was some congruence in the simulation results by species across the models. The
488 share in the recruitment of Silver fir (*Abies alba*) was captured well across the gradient by most
489 models except for xComp, PICUS, and LandClim, which overestimated the recruitment share of
490 this species. There were also some peculiarities evident for some models, such as 4C that did
491 not simulate *Abies alba*. The patterns across models for *Fagus sylvatica* recruitment were more
492 complex, as its recruitment was underestimated by many models across the environmental gradients
493 featuring negligible regeneration at most sites, while others overestimated it in different parts of
494 the environmental space. Most models overestimated the share of *Picea abies* recruitment in the
495 cold-wet part of the gradients. *Pinus sylvestris* and *Quercus* spp. represented a small share of the
496 recruitment in the observed data mostly at warm-dry sites. Many models, however, erroneously
497 featured recruitment for these two species along most of the climatic gradients, although with a
498 low share of basal area. Five models (4C, FORMIND, SIBYLA, xComp, and LandClim) had
499 very small amounts of recruitment of *Pinus sylvestris* and *Quercus* spp., or did not feature any
500 recruitment of these species at all.

501 Lastly, there is no model that performed well across the five species regarding the species-specific
502 recruitment levels in the environmental space (Figure 8). Some models (e.g., ForClim11, PICUS)
503 tended to consistently overestimate the recruitment share of all five species, but most models over-
504 estimated the recruitment share of some species while underestimating the share of others (e.g.,
505 ForCEEPS, LPJ-GUESS).

506 Discussion

507 Tree regeneration is a fundamental process in forest dynamics. Correctly capturing it in dynamic
508 models is fundamental to, for example, evaluate post-disturbance dynamics and potential long-term
509 recovery trajectories as it will define the forest state in the first decades ([Seidl and Turner 2022](#)). If
510 this initial phase is not captured well, we can not properly assess aspects such as forest resilience
511 or the timing, magnitude and progression of carbon sequestration. In this study, for the first time,
512 projections of tree recruitment from multiple models of forest dynamics were confronted with a
513 unique dataset from unmanaged forest reserves across a large environmental gradient in Europe.

514 The EuFoRIa data ([Käber et al. 2023](#)) are exceptional, particularly with respect to the number
515 of records (number of sites and repeated measurements on unmanaged forests), which is essential
516 for capturing a highly ‘noisy’ process such as tree regeneration. The use of this dataset for model
517 benchmarking provided novel insights on the ability of state-of-the-art models to accurately sim-
518 ulate recruitment levels, its species composition and mortality in an early stage of tree life, i.e.,
519 between tree diameters of 7 and 10 cm. Overall, by adopting this approach a much broader un-
520 derstanding resulted than if we had used these data for model calibration: it is primarily from the
521 shortcomings of the models that we can gain ecological insights (cf. [Trugman 2022](#)).

522 Recruitment levels

523 Most of the models overestimated tree recruitment levels. This has potentially far-reaching impli-
524 cations, e.g., regarding biomass (and thus carbon) turnover, with a potential overestimation of the
525 capacity for forest carbon sequestration ([Pugh et al. 2019; Pan et al. 2011](#)). Yet, we focused on one

526 specific stage, i.e., recruitment into the 7 and 10 cm diameter classes. Trees at this size contribute
527 little to carbon sequestration, and if the excess regeneration at this stage is compensated – in the
528 models – soon thereafter by higher mortality, overall simulated tree population dynamics may still
529 be trustworthy at the level of stand structure and attributes such as basal area or biomass.

530 From a modeling point of view, excess tree recruitment inevitably requires excess mortality rates in
531 a later stage, either – as observed for some models – between the 7 and 10 cm diameter thresholds,
532 or soon after the 10 cm threshold has been crossed. In any case, correcting at early stages for the
533 expected forest densities at later stages is equivalent to compensating for a first error (excess re-
534 cruitment) by a second error (excess mortality). It is highly likely that biased projections will result,
535 because the two errors are unlikely to be perfectly linked and thus will not always compensate each
536 other. Hence, this structural problem of most models investigated here is problematic particularly
537 if the models are to be used under novel conditions such as under climate change (e.g., Huber et al.
538 2021) or in a decision support context (e.g., Thriplleton et al. 2021).

539 Recruitment levels define the structure and composition of future forests, but it is equally important
540 to correctly identify areas where regeneration is lacking (Rammer et al. 2021). There are multiple
541 constraints to the regeneration niche of tree species (Price et al. 2001), and therefore the absence
542 of regeneration is likely to be common (Fortin and DeBlois 2007), even over larger areas such as
543 the one-hectare plots used here. Tree regeneration data are characterized by zero-inflation. This
544 was clearly evident from the EuFoRIa dataset, but some of the models did not produce zeros at
545 all, or featured a very low proportion of zero data. This substantial difference may be due to the
546 fact that the simulation results were drawn from equilibrium forests, whereas in reality many of the
547 forest reserves are recovering from past management activities and have become denser over the

548 past decades (e.g., [Heiri et al. 2009](#)), leading to less regeneration than in an equilibrium situation.
549 Another possible reason for this difference may be the exclusion of factors like deer browsing or
550 failure to accurately measure establishment filters, such as competition with the herb layer, site-
551 specific or environmental limitations.

552 Species composition of recruitment

553 Correctly capturing species composition in tree recruitment is important to assess the future func-
554 tional diversity of a forest, e.g., its sensitivity to drought or resilience to disturbances ([Redmond](#)
555 [et al. 2015](#); [Seidl and Turner 2022](#)). In the simulations, overall species diversity levels in the re-
556 cruitment were well within the observed range for half of the models (7 out of 14). Thus, while
557 most models are facing difficulties to quantitatively match recruitment levels (cf. above), their
558 performance is better regarding the composition of recruitment as a function of abiotic and biotic
559 conditions. Most models maintained or even decreased recruitment diversity between the 7 and 10
560 cm thresholds, and the same is visible from the empirical data, but the differences between the 7
561 and 10 cm thresholds were not significant. Diversity variations may be explained by the way the
562 individual models consider regeneration processes ([König et al. 2022](#)), e.g., regeneration can be
563 linked to the seed/seedling bank, seed rain from dispersal or from a pool of available species.

564 The species diversity of the entire stand was captured better by the models than the diversity of
565 tree recruitment, and only a few models overpredicted stand-level diversity. Defining recruitment
566 composition as being proportional to local adult abundance, regardless of productivity, might be
567 a simple and conservative assumption to maintain relative species abundances ([Hanbury-Brown,](#)
568 [Ward, and Kueppers 2022](#)), although this may be simplistic. Furthermore, based on the simulation

569 results, there is no evidence that models with feedback from the canopy (in terms of species com-
570 position of regeneration) captured the species diversity of the recruitment better than those without
571 feedback, with the exception of ForCEEPS that featured significantly more accurate recruitment
572 species diversity with the model version that included feedback (ForCEEPS(f)). The similar per-
573 formance of models with and without feedback is likely because the models put more weight on the
574 regeneration niche arising from abiotic and biotic filters, than from the habitat niche of the adult
575 trees ([Grubb 1977](#)).

576 Recruitment mortality

577 There are several factors that lead to mortality during the regeneration phase, such as competition
578 ([Casper and Jackson 1997](#)) and multiple abiotic factors ([Schmid et al. 2021; Cunningham et al.](#)
579 [2006](#)). As mentioned above, it is reasonable to expect that models that overestimate recruitment at
580 7 cm may have a particularly high mortality between the 7 and 10 cm thresholds. However, this
581 was not consistently evident from the simulations. This implies that these models must have an
582 excess mortality in later stages, if we assume that all models were able to capture the structure and
583 composition of the adult stands along the EuFoRIa gradient; this however was not tested here.

584 Higher mortality towards the adult phase has important implications for forest dynamics and the
585 goods and services provided by forests. On the one hand, mortality in later stages may erroneously
586 enhance the share of less shade tolerant species in the models ([Klopčič, Simončič, and Bončina](#)
587 [2015](#)), therefore shifting the species composition. Unrealistic high stem densities for a longer period
588 of time may overestimate the role of tree regrowth in carbon sink dynamics ([Pugh et al. 2019](#)).
589 Given our set of simulations and analyses, we cannot conclusively assess what is happening in the

590 models, and further studies are required that focus on a wider range of tree sizes and the fate of tree
591 regeneration along such a size continuum.

592 Model performance and model traits

593 In spite of the critical considerations above, it is remarkable that most models did not deviate ex-
594 ceedingly from the observations with respect to simulated recruited levels – after all, few if any
595 of the models' regeneration routines are well constrained by data, with the exception of the em-
596 pirical models xComp and SIBYLA. It is noteworthy that in spite of their empirical basis, these
597 two models did not match empirical recruitment levels, in a similar magnitude as the other models
598 and even in opposite directions (SIBYLA: underestimation; xComp: overestimation). It appears
599 that using empirical data of limited geographical scope that are constrained to managed forests, as
600 done in these models, leads to extrapolation problems already under current climatic conditions.

601 For the other models, whose regeneration modules are not strongly constrained by empirical data,
602 the multiple strategies that are available for modeling regeneration processes (König et al. 2022)
603 could have implied that model performance would be much worse than what we found.

604 Our study showed that increasing complexity in the regeneration modules is not linked to a higher
605 accuracy of the projections of recruitment levels, species composition or mortality at early tree
606 stages, as there was no significant relationship between model performance and model complexity.

607 Increasing complexity in regeneration modules has been called for a long time ago (e.g., Price et al.
608 (2001)) motivated by better process understanding and with the goal of enhancing model accuracy
609 (Bugmann and Seidl 2022). However, more complex models do may not necessarily lead to better
610 projections but rather to reduced transparency and lower predictive power (Franklin et al. 2020).

611 Thus, the question regarding the level of detail that is appropriate and parsimonious for modeling
612 tree regeneration remains open (König et al. 2022; Bugmann and Seidl 2022).

613 Regeneration gradients on regeneration niches

614 Competition for light as a strong filter for tree regeneration has been widely documented (Messier
615 et al. 1999; Collet and Chenost 2006; Berdanier and Clark 2016), but the models examined here
616 did not reproduce this expectation (i.e., decreasing recruitment levels with increasing stand basal
617 area). However, it is difficult to measure light availability at large spatial and temporal scales.
618 We used total stand basal area as a proxy for light availability (cf. Schmid et al. 2021). Yet,
619 we were unable to consider light availability restrictions caused by ground vegetation, which may
620 be an important filter for forest dynamics at least in some EuFoRIa stands (Woltjer et al. 2008).
621 We found pronounced differences in the ranges of stand basal area simulated by the models, but
622 also between models and observations. This made it impossible to evaluate tree recruitment for the
623 extremes of the stand density ranges in some models. This is unfortunate because recruitment levels
624 at low stand density are relevant to assess how well forests are recovering e.g. after gap creation
625 due to disturbance (Seidl and Turner 2022; Grubb 1977). At the level of the simulated one-hectare
626 samples, average basal area is typically not very low as long as no larger disturbances occur, which
627 was explicitly excluded in our simulation protocol. Towards the other end of the spectrum, i.e., with
628 increasing stand basal area, it would be reasonable to expect that the recruitment of the different
629 species would become sparser and drop out entirely under low-light conditions (Klopčič, Poljanec,
630 and Boncina 2012; Zell et al. 2019). However, few models showed this trend, thus indicating
631 that the relation between regeneration and light availability is not yet captured correctly in most

632 models. Yet, several of the models that did feature an increase of regeneration with increasing
633 basal area include a feedback between seed production of mature trees and regeneration. Thus, it
634 seems that in these models higher light competition does not sufficiently compensate for increased
635 seed availability with higher basal area due to a higher abundance of mature trees.

636 Lastly, there were pronounced differences how the main tree species were represented by the models
637 along the environmental gradients, in particular the dominant species *Abies alba*, *Fagus sylvatica*
638 and *Picea abies*. The recruitment levels were sampled from simulations in the equilibrium, and
639 in this state it is expected that non-climax species such as *Pinus sylvestris* or *Quercus* would be
640 of minor importance, or absent ([Klopčič, Simončič, and Bončina 2015](#)). Most models captured
641 this low abundance, which is also found in the empirical data. Thus, although the broad patterns
642 are matched by many models, improvements in the quantification of the regeneration niche of the
643 species are needed, but this cannot be done in the absence of robust datasets across multiple species.

644 Methodological considerations

645 The EuFoRIa data as used here are well suited to better understand tree regeneration. However,
646 three aspects of these data may represent considerable limitations. First, we made a comparison
647 of tree regeneration in an equilibrium state, but we cannot assess how close the forests included
648 in the EuFoRIa dataset were to such a state. The data were collected in forest reserves where no
649 management has taken place for long periods of time. This makes our assumption of an equilibrium
650 between forest properties and environmental drivers more reasonable than it might appear at first
651 sight. In an analysis of primeval forests at demographic equilibrium, Brzeziecki et al. ([2021](#)) found
652 higher recruitment rates than the ones observed in the EuFoRIa data. Thus, the overestimation of

recruitment rates by the models may not be so problematic. Second, the data were collected for recruitment above a 7 cm threshold, thus limiting the assessment of tree regeneration to a specific point of stand dynamics. This constituted a hard limit based on which we can understand only some aspects of tree regeneration, which in its entirety often comprises a rather long period since seed production (Price et al. 2001). In reality, many environmental constraints are acting on young trees (Käber et al. 2021) that we were unable to assess. Yet, recruitment data with lower thresholds are simply not available in a harmonized manner across large environmental gradients. Third, the empirical data were collected from rather small plots, while we sampled simulated recruitment levels from 1 ha areas, which may lead to an incorrect representation of space. Even though the strategy we adopted is not ideal, it represents a common challenge when harmonizing diverse data sources originating from varied sampling strategies (Portier et al. 2022). It would have been extremely challenging for such a variety of models to follow a protocol where the spatial sampling size was different at each of the 200 sites, and it would have introduced additional uncertainty in the results.

The design of our sampling protocol did not include spatial aspects such as seed dispersal or detailed soil data. While we considered a wide range of models of forest dynamics, from stand to global scales, the simulation setup was limited a spatial scale of 1 ha. This lack of consideration of spatial scale relationships is appropriate for stand-scale models, but it potentially puts the landscape models at a disadvantage, as they have been built to be accurate at the landscape level: without any spatial context, we are limiting tree regeneration to the seed influx from the stand itself, unless the model has a background seed input. Yet, the global models should not be at a disadvantage here due to their inherently limited spatial considerations (but cf. Snell et al. 2018; or Lehsten et al. 2019), as they are usually lacking dispersal between cells and are based on a strong abstraction of horizontal space

675 (Hanbury-Brown, Ward, and Kueppers 2022). Lastly, detailed data on soil conditions were not
676 available from the observed data, and independent, admittedly coarse data for soil properties and the
677 climatic water balance had to be used instead. It is noteworthy that many models represent drought
678 using detailed indicators based e.g. on soil water holding capacity, which had to be derived from
679 a rough soil quality measure. This may at least partially explain the unsatisfactory performance of
680 many models along the drought axis (i.e., climatic water balance).

681 Research recommendations

682 With our study, we have demonstrated that models of forest dynamics need a focus on their regen-
683 eration modules to make them more robust. It remains uncertain what level of detail is required
684 to model tree regeneration, and this must be addressed in future research. We recommend that the
685 improvement of the regeneration modules is implemented as additional features that can be traced
686 back, as done here for the variants of ForClim and ForCEEPS, and that model complexity and
687 structure must always be connected with modeling objectives (Albrich et al. 2020). If it should be
688 necessary to include more detail in the regeneration models, this will come with higher parameteri-
689 sation efforts. This will most likely lead to lower generalization because the required data will have
690 to be collected from specific locations, as currently there is no general, comprehensive regeneration
691 dataset available.

692 Therefore, we further recommend that more effort should be invested into collecting harmonized
693 datasets in a site-specific manner covering the different aspects leading to tree regeneration. We
694 emphasize that datasets such as EuFoRla are invaluable and should be expanded in both their spatial
695 extent (e.g., towards boreal and Mediterranean conditions) as well as in time (e.g., continuing the

696 monitoring into the future). Such data will allow for a better evaluation of forest models and help
697 to reduce the uncertainty in their projections, which is crucial when they are used as tools for
698 predicting e.g. the impacts of anthropogenic climate change.

700
701
702
703
704

699 In the present study, we have considered tree regeneration in the equilibrium state only. It is equally
important to understand how these models project tree regeneration after changes in forest struc-
ture by disturbances ([Seidl and Turner 2022](#)), or under different management strategies ([Lindner,](#)
[Lasch, and Erhard 2000](#)). However, this will require an entirely different set of observed data, and
potentially not all forest models would be able to assess the relationship of these aspects on tree
regeneration, e.g. due to the lack of disturbance or appropriate management modules.

705 We recommend to investigate in detail the implications of the current modeling strategies for tree
706 regeneration and, ultimately, simulated forest stand structure. This applies particularly to the er-
707 roneous patterns of excess tree regeneration and later excess mortality, by focusing on a wider
708 range of tree sizes and the related regeneration dynamics. We also recommend, especially for the
709 landscape-level models, the inclusion of explicit spatial considerations regarding tree regeneration
710 ([Beckage and Clark 2003](#)); this, however, is a serious challenge regarding appropriate datasets, as
711 large-scale inventory data have a wide coverage, but by definition do not allow for the assessment
712 of spatial interactions.

713 Exercises like the one presented here, where the models are operated in “blind flight” mode, i.e.,
714 without the possibility of being tuned towards capturing the expected patterns, should be repeated.
715 Such benchmarking exercises should next focus on aspects such as specific model traits and the
716 ecological formulations of particular (sub-)processes, to better understand the implications of the
717 assumptions on which the models are based. Furthermore, the inclusion of a wide range of models

718 with different (1) scales, (2) approaches for capturing population structure, (3) tree regeneration
719 modules and (4) complexity of their formulations will ensure a large benefit to the entire modeling
720 community and beyond.

721 **Conclusions**

722 Models of forest dynamics are important tools in science and decision support, and the formulation
723 of tree regeneration has strong implications for simulated forest properties. The 15 models and
724 variants used here are facing similar challenges in their representation of tree regeneration: they
725 generally overestimate tree recruitment levels, and the simulated regeneration niche is not always
726 captured accurately as a function of biotic (light) and abiotic (temperature and moisture) factors.

727 However, most models properly capture the diversity of the initial tree community, and differences
728 between model formulations, e.g., the presence or absence of feedback from the adult trees did not
729 have a strong effect on capturing the species composition of regeneration.

730 Regarding mortality in the early phase of tree life, many models that feature a particularly high
731 overestimation of recruitment levels are compensating for this by a larger tree mortality. Often,
732 this compensation is not sufficient to reduce the high recruitment levels to realistic values. Overall,
733 there is no clear mortality pattern across all models.

734 When capturing tree regeneration, the specific design decisions taken in the development of any
735 model are more important for its behavior (accuracy) than scale (stand, landscape, global), mod-
736 elling approach (empirical vs. process-based), and complexity. Having both empirical and process-
737 based models in our set, the empirically-based models could have been expected to have a better
738 performance, as they were calibrated with inventory data, but this was not the case. Similarly,

739 higher model complexity does not represent an improvement for capturing tree regeneration.

740 Even though the regeneration routines of most of the models investigated here have never been
741 constrained well by robust data, their projections of recruitment are not overly off. This indicates
742 that a lot can be gained by a focus on the modeling of regeneration processes. The representation
743 of forest dynamics in these models would become much more robust particularly in the face of
744 climate change and post-disturbance dynamics, thus strongly reducing the uncertainty in long-term
745 projections of future forest dynamics.

746 Acknowledgements

747 This paper arises from a workshop held in Davos (Switzerland) in June 2022, co-led by Olalla Díaz-
748 Yáñez, Yannek Käber and Harald Bugmann. Funding for the workshop was provided by COST
749 Action CA19139 PROCLIAS (PROcess-based models for CLimate Impact Attribution across Sec-
750 tors), supported by COST (European Cooperation in Science and Technology; <https://www.cost.eu>)
751 and ETH Zurich (Research Grant ETH-35 18-1). ODY was funded by the Swiss National Science
752 Foundation project 200020_188882 / 1. KB received support from the European Research Coun-
753 cil under the European Union's Horizon 2020 research and innovation program (Grant Agreement
754 101001905). JB acknowledges funding from the long-term research development project no. RVO
755 67985939 from the Czech Academy of Sciences. MM acknowledges financial support from I-
756 Maestro (Innovative forest management strategies for a resilient bioeconomy under climate change
757 and disturbances, grant no. 22035418, 2019–2022) funded by the ERA-NET Cofund Forest-
758 Value. MMina acknowledges funding from the H2020-MSCA-IF project REINFORCE (Grant

759 No. 891671). KM received funding from the Project No. CZ.02.1.01/0.0/0.0/16_019/0000803 fi-
760 nanced by OP RDE, grant no. ITMS2014+313011W580 supported by the Integrated Infrastructure
761 Operational Programme funded by the ERDF. CH acknowledges financial support from Waldfonds
762 WASIM2100 project. We would also like to thank all the people who helped to make this study
763 possible with contributions at different stages particularly Martin Gutsch and Giorgio Vacchiano.

764 Author contributions

765 O. Díaz-Yáñez and Y. Käber contributed equally and share the first authorship. H. Bugmann, Y.
766 Käber, and O. Díaz-Yáñez developed the study design. Y. Käber and O. Díaz-Yáñez prepared the
767 input data used by all modeling teams. All authors contributed by conducting the simulations with
768 their respective models. O. Díaz-Yáñez led the analysis of the simulation results with contributions
769 during the workshop from Y. Käber, T. Anders, K. H. Braziunas, J. Brůna, S. M. Fischer, J. Hetzer,
770 T. Hickler, H. Lischke, M. Mahnken, P. Mairotta, K. Merganičová, T. Mette, X. Morin, W. Rammer,
771 and D. Scherrer. O. Díaz-Yáñez and H. Bugmann led the writing of the manuscript and provided the
772 first draft. All authors participated in the revision of the manuscript and approved its submission.

773 Conflict of interest statement

774 The authors declare no conflicts of interest.

775 References

- 776 Albrich, Katharina, Werner Rammer, Monica G. Turner, Zak Ratajczak, Kristin H. Braziunas,
777 Winslow D. Hansen, and Rupert Seidl. 2020. "Simulating Forest Resilience: A Review."
778 Edited by Thomas Hickler. *Global Ecology and Biogeography* 29 (12): 2082–96. <https://doi.org/10.1111/geb.13197>.
- 780 Allen, Jon C. 1976. "A Modified Sine Wave Method for Calculating Degree Days1." *Environmen-*
781 *tal Entomology* 5 (3): 388–96. <https://doi.org/10.1093/ee/5.3.388>.
- 782 ASTER Science Team. 2019. "ASTER Global Digital Elevation Model V003. 2019, Distributed
783 by NASA EOSDIS Land Processes DAAC, <Https://Doi.org/10.5067/ASTER/ASTGTM.003>.
784 Accessed 2022-11-18." <Https://lpdaac.usgs.gov/products/astgtmv002/>.
- 785 Beckage, Brian, and James S. Clark. 2003. "SEEDLING SURVIVAL AND GROWTH OF THREE
786 FOREST TREE SPECIES: THE ROLE OF SPATIAL HETEROGENEITY." *Ecology* 84 (7):
787 1849–61. [Https://doi.org/10.1890/0012-9658\(2003\)084%5B1849:SSAGOT%5D2.0.CO;2](Https://doi.org/10.1890/0012-9658(2003)084%5B1849:SSAGOT%5D2.0.CO;2).
- 788 Berdanier, Aaron B., and James S. Clark. 2016. "Divergent Reproductive Allocation Trade-Offs
789 with Canopy Exposure Across Tree Species in Temperate Forests." *Ecosphere* 7 (6): e01313.
790 <Https://doi.org/10.1002/ecs2.1313>.
- 791 Bogdziewicz, Michał, Andrew Hacket-Pain, Dave Kelly, Peter A. Thomas, Jonathan Lageard,
792 and Andrew J. Tanentzap. 2021. "Climate Warming Causes Mast Seeding to Break down
793 by Reducing Sensitivity to Weather Cues." *Global Change Biology* 27 (9): 1952–61. <Https://doi.org/10.1111/gcb.15560>.
- 795 Bohn, Friedrich J., Karin Frank, and Andreas Huth. 2014. "Of Climate and Its Resulting Tree
796 Growth: Simulating the Productivity of Temperate Forests." *Ecological Modelling* 278 (April):

- 797 9–17. <https://doi.org/10.1016/j.ecolmodel.2014.01.021>.
- 798 Botkin, Daniel B., James F. Janak, and James R. Wallis. 1972. “Some Ecological Consequences
799 of a Computer Model of Forest Growth.” *The Journal of Ecology* 60 (3): 849. <https://doi.org/10.2307/2258570>.
- 800
- 801 Brzeziecki, Bogdan, Stanisław Drozdowski, Kamil Bielak, Marcin Czacharowski, Jacek
802 Zajączkowski, Włodzimierz Buraczyk, and Leszek Gawron. 2021. “A Demographic
803 Equilibrium Approach to Stocking Control in Mixed, Multiaged Stands in Białowieża
804 Forest, Northeast Poland.” *Forest Ecology and Management* 481 (February): 118694.
805 <https://doi.org/10.1016/j.foreco.2020.118694>.
- 806 Bugmann, Harald. 1996. “A Simplified Forest Model to Study Species Composition Along Cli-
807 mate Gradients.” *Ecology* 77 (7): 2055–74. <https://doi.org/10.2307/2265700>.
- 808 Bugmann, Harald, and Rupert Seidl. 2022. “The Evolution, Complexity and Diversity of Models
809 of Long-Term Forest Dynamics.” *Journal of Ecology* 110 (10): 2288–2307. <https://doi.org/10.1111/1365-2745.13989>.
- 810
- 811 Bugmann, Harald, Rupert Seidl, Florian Hartig, Friedrich Bohn, Josef Brůna, Maxime Cailleret,
812 Louis François, et al. 2019. “Tree Mortality Submodels Drive Simulated Long-Term Forest
813 Dynamics: Assessing 15 Models from the Stand to Global Scale.” *Ecosphere* 10 (2): e02616.
814 <https://doi.org/10.1002/ecs2.2616>.
- 815 Bugmann, Harald, Xiaodong Yan, Martin T. Sykes, Philippe Martin, Marcus Lindner, Paul V. De-
816 sanker, and Steve G. Cumming. 1996. “A Comparison of Forest Gap Models: Model Structure
817 and Behaviour.” *Climatic Change* 34 (2). <https://doi.org/10.1007/BF00224640>.
- 818 Cailleret, Maxime, Steven Jansen, Elisabeth M. R. Robert, Lucía Desoto, Tuomas Aakala, Joseph
819 A. Antos, Barbara Beikircher, et al. 2017. “A Synthesis of Radial Growth Patterns Preceding

- 820 Tree Mortality.” *Global Change Biology* 23 (4): 1675–90. <https://doi.org/10.1111/gcb.13535>.
- 821 Casper, Brenda B., and Robert B. Jackson. 1997. “Plant Competition Underground.” *Annual
Review of Ecology and Systematics* 28 (1): 545–70. <https://doi.org/10.1146/annurev.ecolsys.28.1.545>.
- 822 Clark, J. S., B. Beckage, P. Camill, B. Cleveland, J. HilleRisLambers, J. Lichter, J. McLachlan,
823 J. Mohan, and P. Wyckoff. 1999. “Interpreting Recruitment Limitation in Forests.” *American
Journal of Botany* 86 (1): 1–16. <https://doi.org/10.2307/2656950>.
- 824 Collet, C., and C. Chenost. 2006. “Using Competition and Light Estimates to Predict Diameter and
825 Height Growth of Naturally Regenerated Beech Seedlings Growing Under Changing Canopy
826 Conditions.” *Forestry* 79 (5): 489–502. <https://doi.org/10.1093/forestry/cpl033>.
- 827 Cunningham, Catherine, Niklaus E Zimmermann, Veronika Stoeckli, and Harald Bugmann. 2006.
828 “Growth Response of Norway Spruce Saplings in Two Forest Gaps in the Swiss Alps to Ar-
829ificial Browsing, Infection with Black Snow Mold, and Competition by Ground Vegetation.”
830 *Canadian Journal of Forest Research* 36 (11): 2782–93. <https://doi.org/10.1139/x06-156>.
- 831 Díaz-Yáñez, Olalla, Yannek Käber, and Harald Bugmann. 2022. “Regeneration Workshop Proto-
832 col.” OSF. <https://doi.org/10.17605/OSF.IO/CZDXP>.
- 833 EuFoRia. 2019. “European Forest Reserves Initiative. Link: <Https://Www.wsl.ch/de/Projekte/Euforia.html#tabelo>
834 Tab2.”
- 835 Fabrika, M. 2005. “Simulátor Biodynamiky Lesa SIBYLA, Konцепcia, Konštrukcia a Programové
836 Riešenie.” *Zvolen, Slovakia, Habilitačná Práca, Technická Univerzita Vo Zvolene*.
- 837 Fabrika, M, and J Ďurský. 2012. “Algorithms and Software Solution of Thinning Models for
838 SIBYLA Growth Simulator.” *Journal of Forest Science* 51 (No. 10): 431–45. <https://doi.org/10.17221/4577-JFS>.

- 843 Fischer, Rico, Friedrich Bohn, Mateus Dantas de Paula, Claudia Dislich, Jürgen Groeneveld, Al-
844 varo G. Gutiérrez, Martin Kazmierczak, et al. 2016. "Lessons Learned from Applying a Forest
845 Gap Model to Understand Ecosystem and Carbon Dynamics of Complex Tropical Forests."
846 *Ecological Modelling* 326 (April): 124–33. <https://doi.org/10.1016/j.ecolmodel.2015.11.018>.
- 847 Fischlin, Andreas, Harald Bugmann, and Dimitrios Gyalistras. 1995. "Sensitivity of a Forest
848 Ecosystem Model to Climate Parametrization Schemes." *Environmental Pollution* 87 (3): 267–
849 82. [https://doi.org/10.1016/0269-7491\(94\)P4158-K](https://doi.org/10.1016/0269-7491(94)P4158-K).
- 850 FOREST EUROPE. 2020. "State of Europe's Forests 2020." Ministerial Conference on the Pro-
851 tection of Forests in Europe - FOREST EUROPE.
- 852 Fortin, Mathieu, and Josianne DeBlois. 2007. "Modeling Tree Recruitment with Zero-Inflated
853 Models: The Example of Hardwood Stands in Southern Québec, Canada." *Forest Sciences* 53
854 (4): 529–39. <https://doi.org/10.1093/forestscience/53.4.529>.
- 855 Franklin, Oskar, Sandy P. Harrison, Roderick Dewar, Caroline E. Farrior, Åke Bränström, Ulf
856 Dieckmann, Stephan Pietsch, et al. 2020. "Organizing Principles for Vegetation Dynamics."
857 *Nature Plants* 6 (5): 444–53. <https://doi.org/10.1038/s41477-020-0655-x>.
- 858 Grubb, P. J. 1977. "THE MAINTENANCE OF SPECIES-RICHNESS IN PLANT COMMUNI-
859 TIES: THE IMPORTANCE OF THE REGENERATION NICHE." *Biological Reviews* 52 (1):
860 107–45. <https://doi.org/10.1111/j.1469-185X.1977.tb01347.x>.
- 861 Gutiérrez, Alvaro G., Rebecca S. Snell, and Harald Bugmann. 2016. "Using a Dynamic Forest
862 Model to Predict Tree Species Distributions: Predicting Tree Species Distribution." *Global
863 Ecology and Biogeography* 25 (3): 347–58. <https://doi.org/10.1111/geb.12421>.
- 864 Hallsby, Göran, Kristina Ahnlund Ulvcrona, Anders Karlsson, Björn Elfving, Hans Sjögren,
865 Thomas Ulvcrona, and Urban Bergsten. 2015. "Effects of Intensity of Forest Regeneration

- 866 Measures on Stand Development in a Nationwide Swedish Field Experiment.” *Forestry* 88
867 (4): 441–53. <https://doi.org/10.1093/forestry/cpv010>.
- 868 Hanbury-Brown, Adam R., Rachel E. Ward, and Lara M. Kueppers. 2022. “Forest Regenera-
869 tion Within Earth System Models: Current Process Representations and Ways Forward.” *New
870 Phytologist* 235 (1): 20–40. <https://doi.org/10.1111/nph.18131>.
- 871 Heiri, C., A. Wolf, L. Rohrer, and H. Bugmann. 2009. “Forty Years of Natural Dynamics in Swiss
872 Beech Forests: Structure, Composition, and the Influence of Former Management.” *Ecological
873 Applications* 19 (7): 1920–34. <https://doi.org/10.1890/08-0516.1>.
- 874 Hengl, Tomislav, Jorge Mendes de Jesus, Gerard B. M. Heuvelink, Maria Ruiperez Gonzalez, Milan
875 Kilibarda, Aleksandar Blagotić, Wei Shangguan, et al. 2017. “SoilGrids250m: Global Gridded
876 Soil Information Based on Machine Learning.” Edited by Ben Bond-Lamberty. *PLOS ONE* 12
877 (2): e0169748. <https://doi.org/10.1371/journal.pone.0169748>.
- 878 Hickler, Thomas, Katrin Vohland, Jane Feehan, Paul A. Miller, Benjamin Smith, Luis Costa,
879 Thomas Giesecke, et al. 2012. “Projecting the Future Distribution of European Potential Nat-
880 ural Vegetation Zones with a Generalized, Tree Species-Based Dynamic Vegetation Model:
881 Future Changes in European Vegetation Zones.” *Global Ecology and Biogeography* 21 (1):
882 50–63. <https://doi.org/10.1111/j.1466-8238.2010.00613.x>.
- 883 Huber, Nica, Harald Bugmann, Maxime Cailleret, Nicolas Bircher, and Valentine Lafond. 2021.
884 “Stand-Scale Climate Change Impacts on Forests over Large Areas: Transient Responses and
885 Projection Uncertainties.” *Ecological Applications* 31 (4). <https://doi.org/10.1002/eap.2313>.
- 886 Huber, Nica, Harald Bugmann, and Valentine Lafond. 2020. “Capturing Ecological Processes in
887 Dynamic Forest Models: Why There Is No Silver Bullet to Cope with Complexity.” *Ecosphere*
888 11 (5). <https://doi.org/10.1002/ecs2.3109>.

- 889 Käber, Yannek, Christof Bigler,..., and Harald Bugmann. 2023. "Sheltered and Suppressed: Tree
890 Regeneration in Unmanaged European Forests (Under Revision)." *Journal of Ecology*, De-
891 cember.
- 892 Käber, Yannek, Peter Meyer, Jonas Stillhard, Emiel De Lombaerde, Jürgen Zell, Golo Stadelmann,
893 Harald Bugmann, and Christof Bigler. 2021. "Tree Recruitment Is Determined by Stand Struc-
894 ture and Shade Tolerance with Uncertain Role of Climate and Water Relations." *Ecology and*
895 *Evolution* 11 (17): 12182–203. <https://doi.org/10.1002/ece3.7984>.
- 896 Karger, Dirk N., Stefan Lange, Chantal Hari, Christopher P. O. Reyer, and Niklaus E. Zimmer-
897 mann. 2021. "CHELSA-W5E5 V1.0: W5E5 V1.0 Downscaled with CHELSA V2.0. ISIMIP
898 Repository. <Https://Doi.org/10.48364/ISIMIP.836809>."
- 899 Klopčič, Matija, Ales Poljanec, and Andrej Boncina. 2012. "Modelling Natural Recruitment of
900 European Beech (*Fagus Sylvatica L.*)."*Forest Ecology and Management* 284 (November):
901 142–51. <https://doi.org/10.1016/j.foreco.2012.07.049>.
- 902 Klopčič, Matija, Tina Simončič, and Andrej Bončina. 2015. "Comparison of Regeneration and
903 Recruitment of Shade-Tolerant and Light-Demanding Tree Species in Mixed Uneven-Aged
904 Forests: Experiences from the Dinaric Region."*Forestry* 88 (5): 552–63. <https://doi.org/10.1093/forestry/cpv021>.
- 905 906 König, Louis A., Frits Mohren, Mart-Jan Schelhaas, Harald Bugmann, and Gert-Jan Nabuurs.
907 2022. "Tree Regeneration in Models of Forest Dynamics – Suitability to Assess Climate
908 Change Impacts on European Forests."*Forest Ecology and Management* 520 (September):
909 120390. <https://doi.org/10.1016/j.foreco.2022.120390>.
- 910 911 Kroiss, Steve J., and Janneke HilleRisLambers. 2015. "Recruitment Limitation of Long-Lived
Conifers: Implications for Climate Change Responses."*Ecology* 96 (5): 1286–97. <https://doi.org/10.1002/ecology.2015.1286>.

- 912 [org/10.1890/14-0595.1](https://doi.org/10.1890/14-0595.1).
- 913 Langan, Liam, Steven I. Higgins, and Simon Scheiter. 2017. “Climate-Biomes, Pedo-Biomes or
914 Pyro-Biomes: Which World View Explains the Tropical Forest-Savanna Boundary in South
915 America?” *Journal of Biogeography* 44 (10): 2319–30. <https://doi.org/10.1111/jbi.13018>.
- 916 Lasch-Born, Petra, Felicitas Suckow, Christopher P. O. Reyer, Martin Gutsch, Chris Kollas, Franz-
917 Werner Badeck, Harald K. M. Bugmann, et al. 2020. “Description and Evaluation of the
918 Process-Based Forest Model 4C V2.2 at Four European Forest Sites.” *Geoscientific Model
919 Development* 13 (11): 5311–43. <https://doi.org/10.5194/gmd-13-5311-2020>.
- 920 Lehsten, Veiko, Michael Mischirow, Erik Lindström, Dörte Lehsten, and Heike Lischke. 2019.
921 “LPJ-GM 1.0: Simulating Migration Efficiently in a Dynamic Vegetation Model.” *Geoscienc-
922 tific Model Development* 12 (3): 893–908. <https://doi.org/10.5194/gmd-12-893-2019>.
- 923 Leishman, Mr, L Hughes, K French, D Armstrong, and M Westoby. 1992. “Seed and Seedling
924 Biology in Relation to Modelling Vegetation Dynamics Under Global Climate Change.” *Aus-
925 tralian Journal of Botany* 40 (5): 599. <https://doi.org/10.1071/BT9920599>.
- 926 Lexer, Manfred J, and Karl Hönninger. 2001. “A Modified 3D-Patch Model for Spatially Explicit
927 Simulation of Vegetation Composition in Heterogeneous Landscapes.” *Forest Ecology and
928 Management*, 23.
- 929 Lindner, Marcus, Petra Lasch, and Markus Erhard. 2000. “Alternative Forest Management Strate-
930 gies Under Climatic Change – Prospects for Gap Model Applications in Risk Analyses.” *Silva
931 Fennica* 34 (2). <https://doi.org/10.14214/sf.634>.
- 932 Lindner, Marcus, Michael Maroschek, Sigrid Netherer, Antoine Kremer, Anna Barbati, Jordi
933 Garcia-Gonzalo, Rupert Seidl, et al. 2010. “Climate Change Impacts, Adaptive Capacity, and
934 Vulnerability of European Forest Ecosystems.” *Forest Ecology and Management* 259 (4):

- 935 698–709. <https://doi.org/10.1016/j.foreco.2009.09.023>.
- 936 Lischke, Heike, and Thomas J. Löffler. 2006. “Intra-Specific Density Dependence Is Required to
937 Maintain Species Diversity in Spatio-Temporal Forest Simulations with Reproduction.” *Eco-*
938 *logical Modelling* 198 (3-4): 341–61. <https://doi.org/10.1016/j.ecolmodel.2006.05.005>.
- 939 Meier, Eliane S., Thomas C. Edwards Jr, Felix Kienast, Matthias Dobbertin, and Niklaus E. Zim-
940 mermann. 2011. “Co-Occurrence Patterns of Trees Along Macro-Climatic Gradients and Their
941 Potential Influence on the Present and Future Distribution of *Fagus Sylvatica* L.: Influence
942 of Co-Occurrence Patterns on *Fagus Sylvatica*.” *Journal of Biogeography* 38 (2): 371–82.
943 <https://doi.org/10.1111/j.1365-2699.2010.02405.x>.
- 944 Messier, Christian, René Doucet, Jean-Claude Ruel, Yves Claveau, Colin Kelly, and Martin J Le-
945 chowicz. 1999. “Functional Ecology of Advance Regeneration in Relation to Light in Boreal
946 Forests” 29: 12.
- 947 Mette, Tobias. 2014. “Modelling Patagonian Lenga-Forest Dynamics (*Nothofagus Pumilio*) in
948 Chile. Final Report, DFG Project ME 3568/2-1, Bonn, Germany.”
- 949 Miina, Jari, Kalle Eerikäinen, and Hubert Hasenauer. 2006. “Modeling Forest Regeneration. In:
950 Hasenauer, h. (Eds.), Sustainable Forest Management. Growth Models for Europe (Pp. 93–
951 110).” In.
- 952 Morin, Xavier, Harald Bugmann, François Coligny, Nicolas Martin-StPaul, Maxime Cailleret,
953 Jean-Marc Limousin, Jean-Marc Ourcival, et al. 2021. “Beyond Forest Succession: A Gap
954 Model to Study Ecosystem Functioning and Tree Community Composition Under Climate
955 Change.” Edited by Julia Koricheva. *Functional Ecology* 35 (4): 955–75. <https://doi.org/10.1111/1365-2435.13760>.
- 956 Muñoz Sabater, J. 2019. “ERA5-Land Hourly Data from 1981 to Present. Copernicus Climate

- 958 Change Service (C3S) Climate Data Store (CDS). (23.01.2022)." [doi:10.24381/cds.e2161bac](https://doi.org/10.24381/cds.e2161bac).
- 959 Pan, Yude, Richard A. Birdsey, Jingyun Fang, Richard Houghton, Pekka E. Kauppi, Werner A.
- 960 Kurz, Oliver L. Phillips, et al. 2011. "A Large and Persistent Carbon Sink in the World's
- 961 Forests." *Science* 333 (6045): 988–93. <https://doi.org/10.1126/science.1201609>.
- 962 Portier, Jeanne, Florian Zellweger, Jürgen Zell, Iciar Alberdi Asensio, Michal Bosela, Johannes
- 963 Breidenbach, Vladimír Šeben, Rafael O. Wüest, and Brigitte Rohner. 2022. "Plot Size Matters:
- 964 Toward Comparable Species Richness Estimates Across Plot-Based Inventories." *Ecology and*
- 965 *Evolution* 12 (6). <https://doi.org/10.1002/ece3.8965>.
- 966 Pretzsch, Hans, and Peter Biber. 2003. "A Re-Evaluation of Reineke's Rule and Stand Density
- 967 Index." *Forest Sciences* 51 (4): 304–20. <https://doi.org/10.1093/forestscience/51.4.304>.
- 968 Pretzsch, H., P. Biber, and J. Durský. 2002. "The Single Tree-Based Stand Simulator SILVA:
- 969 Construction, Application and Evaluation." *Forest Ecology and Management* 162 (1): 3–21.
- 970 [https://doi.org/10.1016/S0378-1127\(02\)00047-6](https://doi.org/10.1016/S0378-1127(02)00047-6).
- 971 Price, David T, Niklaus E Zimmermann, Manfred J Lexer, Paul Leadley, Irma T M Jorritsma,
- 972 Jörg Schaber, Donald F Clark, et al. 2001. "Regeneration in Gap Models: Priority Issues for
- 973 Studying Forest Responses to Climate Change" 51: 34.
- 974 Pugh, Thomas A. M., Mats Lindeskog, Benjamin Smith, Benjamin Poulter, Almut Arneth, Vanessa
- 975 Haverd, and Leonardo Calle. 2019. "Role of Forest Regrowth in Global Carbon Sink Dynam-
- 976 ics." *Proceedings of the National Academy of Sciences* 116 (10): 4382–87. <https://doi.org/10.1073/pnas.1810512116>.
- 977 Rammer, Werner, Kristin H. Braziunas, Winslow D. Hansen, Zak Ratajczak, Anthony L. Wester-
- 979 ling, Monica G. Turner, and Rupert Seidl. 2021. "Widespread Regeneration Failure in Forests
- 980 of Greater Yellowstone Under Scenarios of Future Climate and Fire." *Global Change Biology*

- 981 27 (18): 4339–51. <https://doi.org/10.1111/gcb.15726>.
- 982 Redmond, Miranda D., Neil S. Cobb, Michael J. Clifford, and Nichole N. Barger. 2015. “Woodland
983 Recovery Following Drought-Induced Tree Mortality Across an Environmental Stress Gradi-
984 ent.” *Global Change Biology* 21 (10): 3685–95. <https://doi.org/10.1111/gcb.12976>.
- 985 Reineke, L H. 1933. “PERFECTING A STAND-DENSITY INDEX FOR EVEN- AGED
986 FORESTS’,” 12.
- 987 Reyer, Christopher, Petra Lasch-Born, Felicitas Suckow, Martin Gutsch, Aline Murawski, and
988 Tobias Pilz. 2014. “Projections of Regional Changes in Forest Net Primary Productivity for
989 Different Tree Species in Europe Driven by Climate Change and Carbon Dioxide.”
- 990 Ribbens, Eric, John A. Silander, and Stephen W. Pacala. 1994. “Seedling Recruitment in Forests:
991 Calibrating Models to Predict Patterns of Tree Seedling Dispersion.” *Ecology* 75 (6): 1794–
992 1806. <https://doi.org/10.2307/1939638>.
- 993 Scheiter, Simon, and S. I. Higgins. 2009. “Impacts of Climate Change on the Vegetation of Africa:
994 An Adaptive Dynamic Vegetation Modelling Approach.” *Global Change Biology*.
- 995 Scheiter, Simon, Liam Langan, and Steven I. Higgins. 2013. “Next-Generation Dynamic Global
996 Vegetation Models: Learning from Community Ecology.” *New Phytologist* 198 (3): 957–69.
997 <https://doi.org/10.1111/nph.12210>.
- 998 Scheller, Robert M., James B. Domingo, Brian R. Sturtevant, Jeremy S. Williams, Arnold Rudy,
999 Eric J. Gustafson, and David J. Mladenoff. 2007. “Design, Development, and Application of
1000 LANDIS-II, a Spatial Landscape Simulation Model with Flexible Temporal and Spatial Res-
1001 olution.” *Ecological Modelling* 201 (3-4): 409–19. <https://doi.org/10.1016/j.ecolmodel.2006.10.009>.
- 1002 [10.1016/j.ecolmodel.2006.10.009](https://doi.org/10.1016/j.ecolmodel.2006.10.009).
- 1003 Schmid, Ueli, Christof Bigler, Monika Frehner, and Harald Bugmann. 2021. “Abiotic and Biotic

- 1004 Determinants of Height Growth of *Picea Abies* Regeneration in Small Forest Gaps in the Swiss
1005 Alps." *Forest Ecology and Management* 490 (June): 119076. <https://doi.org/10.1016/j.foreco.2021.119076>.
- 1006
- 1007 Schumacher, Sabine, Björn Reineking, Jason Sibold, and Harald Bugmann. 2006. "Modeling
1008 the Impact of Climate and Vegetation on Fire Regimes in Mountain Landscapes." *Landscape
1009 Ecology* 21 (4): 539–54. <https://doi.org/10.1007/s10980-005-2165-7>.
- 1010 Seidl, Rupert, Werner Rammer, Robert M. Scheller, and Thomas A. Spies. 2012. "An Individual-
1011 Based Process Model to Simulate Landscape-Scale Forest Ecosystem Dynamics." *Ecological
1012 Modelling* 231 (April): 87–100. <https://doi.org/10.1016/j.ecolmodel.2012.02.015>.
- 1013 Seidl, Rupert, and Monica G. Turner. 2022. "Post-Disturbance Reorganization of Forest Ecosys-
1014 tems in a Changing World." *Proceedings of the National Academy of Sciences* 119 (28):
1015 e2202190119. <https://doi.org/10.1073/pnas.2202190119>.
- 1016 Shugart, Herman Henry. 1984. *A Theory of Forest Dynamics. The Ecological Implications of
1017 Forest Succession Models*. NY, USA: Springer-Verlag.
- 1018 Smith, Benjamin, I. Colin Prentice, and Martin T. Sykes. 2001a. "Representation of Vegetation
1019 Dynamics in the Modelling of Terrestrial Ecosystems: Comparing Two Contrasting Approaches
1020 Within European Climate Space: *Vegetation Dynamics in Ecosystem Models*." *Global Ecology
1021 and Biogeography* 10 (6): 621–37. <https://doi.org/10.1046/j.1466-822X.2001.t01-1-00256.x>.
- 1022 ———. 2001b. "Representation of Vegetation Dynamics in the Modelling of Terrestrial Ecosys-
1023 tems: Comparing Two Contrasting Approaches Within European Climate Space: *Vegetation
1024 Dynamics in Ecosystem Models*." *Global Ecology and Biogeography* 10 (6): 621–37. <https://doi.org/10.1046/j.1466-822X.2001.t01-1-00256.x>.
- 1025
- 1026 Snell, Rebecca S., Ché Elkin, Sven Kotlarski, and Harald Bugmann. 2018. "Importance of Climate

- 1027 Uncertainty for Projections of Forest Ecosystem Services.” *Regional Environmental Change*
1028 18 (7): 2145–59. <https://doi.org/10.1007/s10113-018-1337-3>.
- 1029 Speich, Mjr. 2019. “Quantifying and Modeling Water Availability in Temperate Forests: A Review
1030 of Drought and Aridity Indices.” *iForest - Biogeosciences and Forestry* 12 (1): 1–16. <https://doi.org/10.3832/ifor2934-011>.
- 1032 Thrippleton, Timothy, Clemens Blattert, Leo Gallus Bont, Reinhard Mey, Jürgen Zell, Esther
1033 Thürig, and Janine Schweier. 2021. “A Multi-Criteria Decision Support System for Strategic
1034 Planning at the Swiss Forest Enterprise Level: Coping With Climate Change and Shifting
1035 Demands in Ecosystem Service Provisioning.” *Frontiers in Forests and Global Change* 4 (August):
1036 693020. <https://doi.org/10.3389/ffgc.2021.693020>.
- 1037 Trugman, Anna T. 2022. “Integrating Plant Physiology and Community Ecology Across Scales
1038 Through Trait-Based Models to Predict Drought Mortality.” *New Phytologist* 234 (1): 21–27.
1039 <https://doi.org/10.1111/nph.17821>.
- 1040 Vacchiano, Giorgio, Davide Ascoli, Fabio Berzaghi, Manuel Esteban Lucas-Borja, Thomas Caig-
1041 nard, Alessio Collalti, Paola Mairotta, et al. 2018. “Reproducing Reproduction: How to Sim-
1042 ulate Mast Seeding in Forest Models.” *Ecological Modelling* 376 (May): 40–53. <https://doi.org/10.1016/j.ecolmodel.2018.03.004>.
- 1044 Vanclay, J. K., and K. Jerome. 1994. *Modelling Forest Growth and Yield: Applications to Mixed
1045 Tropical Forests*. Wallingford, U.K: CAB International.
- 1046 Vanclay, J. K., and J. P. Skovsgaard. 1997. “Evaluating Forest Growth Models.” *Ecological Mod-
1047 elling* 98 (1): 1–12. [https://doi.org/10.1016/S0304-3800\(96\)01932-1](https://doi.org/10.1016/S0304-3800(96)01932-1).
- 1048 Walck, Jeffrey L., Siti N. Hidayati, Kingsley W. Dixon, Ken Thompson, and Peter Poschlod. 2011.
1049 “Climate Change and Plant Regeneration from Seed.” *Global Change Biology* 17 (6): 2145–61.

- 1050 <https://doi.org/10.1111/j.1365-2486.2010.02368.x>.
- 1051 Wehrli, A., P. J. Weisberg, W. Schönenberger, P. Brang, and H. Bugmann. 2006. "Improving the
1052 Establishment Submodel of a Forest Patch Model to Assess the Long-Term Protective Effect
1053 of Mountain Forests." *European Journal of Forest Research* 126 (1): 131–45. <https://doi.org/10.1007/s10342-006-0142-6>.
- 1054
- 1055 Woltjer, M., W. Rammer, M. Brauner, R. Seidl, G. M. J. Mohren, and M. J. Lexer. 2008. "Cou-
1056 pling a 3D Patch Model and a Rockfall Module to Assess Rockfall Protection in Mountain
1057 Forests." *Journal of Environmental Management* 87 (3): 373–88. <https://doi.org/10.1016/j.jenvman.2007.01.031>.
- 1058
- 1059 Wood, S N. 2011. "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estima-
1060 tion of Semiparametric Generalized Linear Models." *Journal of the Royal Statistical Society*
1061 73 (1): 3–36. <https://doi.org/doi.org/10.1111/j.1467-9868.2010.00749.x>.
- 1062 Zell, Jürgen, Brigitte Rohner, Esther Thürig, and Golo Stadelmann. 2019. "Modeling Ingrowth
1063 for Empirical Forest Prediction Systems." *Forest Ecology and Management* 433 (February):
1064 771–79. <https://doi.org/10.1016/j.foreco.2018.11.052>.

1065 **Table**

1066 **Table 1.** Models included in this study, their characteristics and simulation strategies used. Scale
1067 refers to the models scale, Type to the models type, Population structure to the model population
1068 approach, Feedback to the inclusion or not of adult tree feedback to the level and species composi-
1069 tion of regeneration, Approach to the model regeneration module approach, Start from to the stage
1070 from which trees are recruited, Complexity, to the mean regeneration formulation as estimated in
1071 (Bugmann & Seidl, 2022), Species to which species from the original eleven species were simu-
1072 lated, Runtime (spin up) [years] to the number of years used per model in their simulations for the
1073 spin up period and regeneration sampling; Climate data type (use of years) to the type of data used
1074 in the simulations and how it was used, and Reference provides the link to the publication with
1075 more model details. Further information can be found on the individual model reports available in
1076 the Appendix S1: Model reports.

Table 1

Model	Scale(a)	Type(b)	Population structure(c)	Feedback(d)	Approach(e)	Start from	Mean regeneration formulation complexity (Bugmann & Seidl 2022)	Species simulated	Runtime for sampling (spin up) [years]	Climate data type (use of years)(f)	Reference
4C	S	PB	C	N	R	saplings	-0.04	F. sylvatica, P. sylvestris, P. abies, Betula spp., Quercus spp.	2000 (500)	D (loop)	Lasch-Born et al. (2020)
ForClim	S	PB	C	N	Rc	ingrowth	0.28 (ForClim 4.0.1)	All	200 (1000)	M (weather generator)	Bugmann (1996)
ForCEEPS	S	PB	I	N / Y (f)	Rc	saplings	0.43	All	2000 (2000)	M (random)	Morin et al. (2021)
FORMIND	S	PB	I	N	Rc	ingrowth	-0.57	All	12000 (1000)	D (random)	Bohn et al. (2014) and Fischer et al. (2016)
PICUS	S	PB	I	Y	R	seeds	0.81	All	2000 (600)	M (random)	Lexer & Hönninger (2001)
SIBYLA	S	EM	I	Y	R	seeds	NA	All	2500 (500)	M (weather generator)	Fabrika (2005)
xComp	S	EM	I	Y	R	seedlings	NA	All (-F. excelsior, -A. glutinosa)	2000 (750)	M (random)	Mette (2014)
iLand	L	PB	I	Y	R	seeds	1.08	All	510 (1500)	D (random)	Seidl et al. (2012)
LandClim	L	PB	C	Y	Rc	ingrowth	0.5	All	2000 (2000)	M (random)	Schumacher et al. (2006)
Landis-II	L	PB	C	Y	Rc	saplings	-0.1	All	1280 (700)	M (random)	Scheller et al. (2007)
TreeMig	L	PB	C	Y	R	seedbank	0.12	All	100 (900)	M (random)	Lischke et al. (2006)
LPJ-GUESS	G	PB	C	Y	Rc	saplings	0.35	All (-A. pseuoplatanus, -A. glutinosa)	37 (500)	D (loop)	Smith et al. (2001) and Hickler et al. (2012)
aDGVM2	G	PB	I	Y	Rc	seedbank	-0.08	Ecological strategies	37 (600)	D (loop)	Langan et al. (2017) and Scheiter et al. (2013)

^a S-Stand, L-Landscape, G-Global^b PB-Process based, EM-Empirical^c C-Cohort, I-Individual^d Y-Yes, N-No^e R-Regeneration, Rc-Recruitment^f D-Daily, M-Monthly

1077 **Figure captions**

1078 **Figure 1.** The 200 sites of the EuFoRia network (EuFoRia, 2019) used in the present study, the
1079 colors show the dominant tree species per site in the observed regeneration at the 10 cm threshold.

1080 **Figure 2.** Mean regeneration levels across all samples per site, plotted for the 200 sites and for each
1081 model. The red dashed lines shows the 25th and 75th percentiles for the 7 cm diameter threshold in
1082 the observed data. There are two boxplots for each model where the left lighter boxplot corresponds
1083 to 7 cm and the right darker boxplot to 10 cm

1084 **Figure 3.** Shannon index of tree regeneration, calculated by basal area, as the mean value across
1085 the 200 samples per site for the diameter thresholds of 7 and 10 cm, respectively. The red dashed
1086 lines shows the 25th and 75th percentiles for the 7 cm diameter threshold in the observed data.

1087 **Figure 4.** Mean Shannon index across all samples per site for observed and simulated data. Each
1088 plot shows one pattern represented by one exemplary model of each category (overpredicted, inter-
1089 mediate and underpredicted). The full data with the grouping of the models are shown in Figures
1090 S3 and S4. n indicates the number of models falling in each group. A) Examples for the three
1091 trends across models for species diversity at the stand level. B) Examples for the three trends
1092 across models for regeneration at the 7 cm threshold (165 sites).

1093 **Figure 5.** Ratio of tree regeneration rates between the 7 and 10 cm thresholds. Dashed blue lines
1094 mark a ratio equal to 1, indicating no decrease in tree regeneration between 7 and 10 cm, and a ratio
1095 equal to 1.77, corresponding to the Reineke self-thinning ratio under even-aged conditions.

1096 **Figure 6.** Ratio of tree regeneration between the 7 and 10 cm thresholds (regeneration 7 cm /
1097 regeneration 10 cm) and the overestimation proportion at 7 cm ([simulated - observed] / observed)

1098 for the mean regeneration per model across sites and samples. The horizontal dashed lines show a
1099 ratio equal to 1, indicating no decrease of regeneration between 7 and 10 cm, and a ratio equal to
1100 1.77 corresponding to the Reineke self-thinning line under even-aged conditions. The vertical line
1101 indicates zero overestimation.

1102 **Figure 7.** Mean recruitment across the 200 samples per site with the y-axis scaled differently by
1103 model, for the 200 sites against gradients of (a) total basal area; (b) climatic water balance; (c)
1104 seasonal degree-day sum. The values were split into ten bins; the red lines represent a GAM model
1105 showing the trend in the observed data. Y-axis are not equally scaled.

1106 **Figure 8.** Regeneration levels of the five main species in terms of their basal area share in the ob-
1107 servations (top row) and the performance of each model across the environmental gradients (other
1108 rows) for the 7 cm regeneration threshold. The values shown are the mean of the 200 samples per
1109 site and across the sites in each bin (tile), with ten bins per gradient. The sizes of the circles repre-
1110 sent the ratio between the regeneration basal area of the species and the total regeneration basal area
1111 for all species. The absence of a circle indicates a zero basal area share in the regeneration, or the
1112 absence of regeneration altogether. The color gradient (for the models only) shows the difference
1113 between the simulated and observed ratio of regeneration basal area of the species and the total
1114 regeneration basal area for all the species.

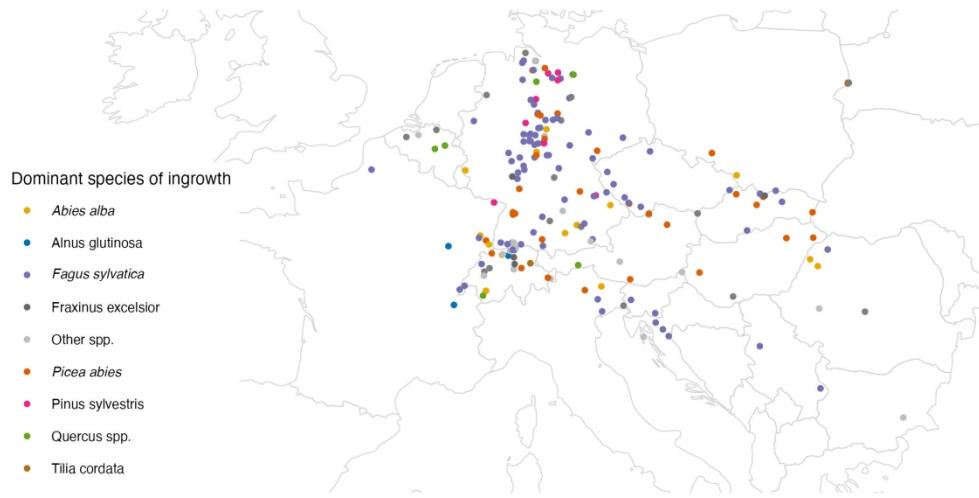


Figure 1. The 200 sites of the EuFoRIa network (EuFoRIa, 2019) used in the present study, the colors show the dominant tree species per site in the observed regeneration at the 10 cm threshold.

787x449mm (72 x 72 DPI)

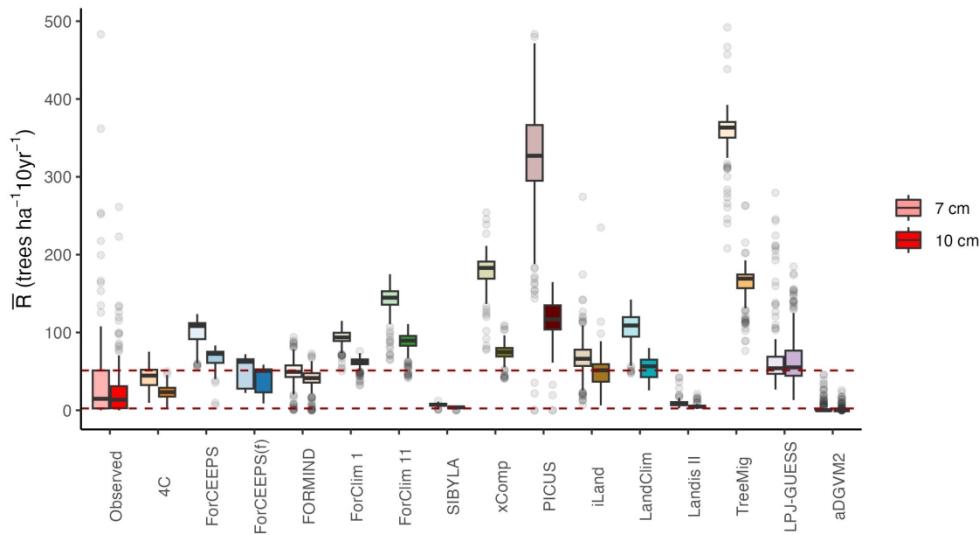


Figure 2. Mean regeneration levels across all samples per site, plotted for the 200 sites and for each model. The red dashed lines shows the 25th and 75th percentiles for the 7 cm diameter threshold in the observed data. There are two boxplots for each model where the left lighter boxplot corresponds to 7 cm and the right darker boxplot to 10 cm

787x449mm (72 x 72 DPI)

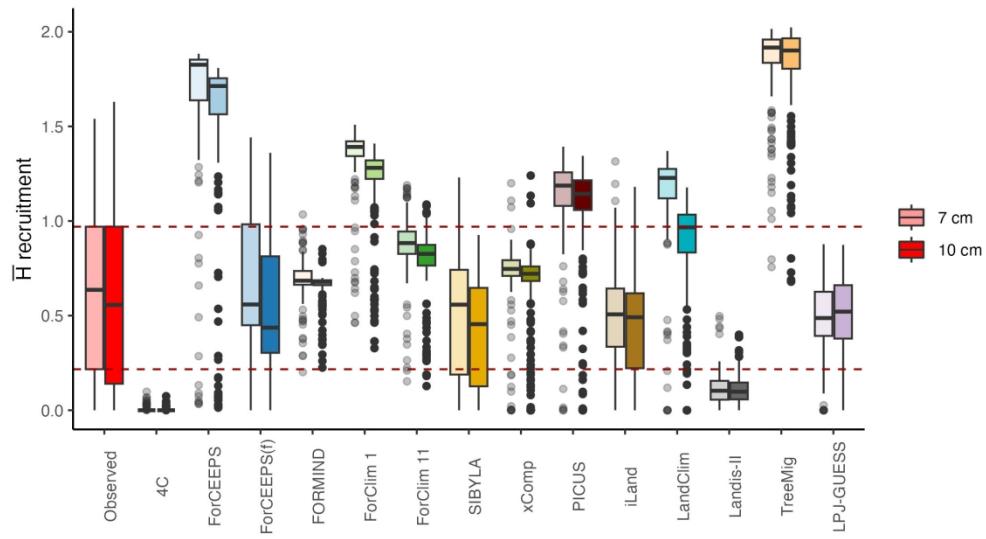


Figure 3. Shannon index of tree regeneration, calculated by basal area, as the mean value across the 200 samples per site for the diameter thresholds of 7 and 10 cm, respectively. The red dashed lines shows the 25th and 75th percentiles for the 7 cm diameter threshold in the observed data.

787x449mm (72 x 72 DPI)

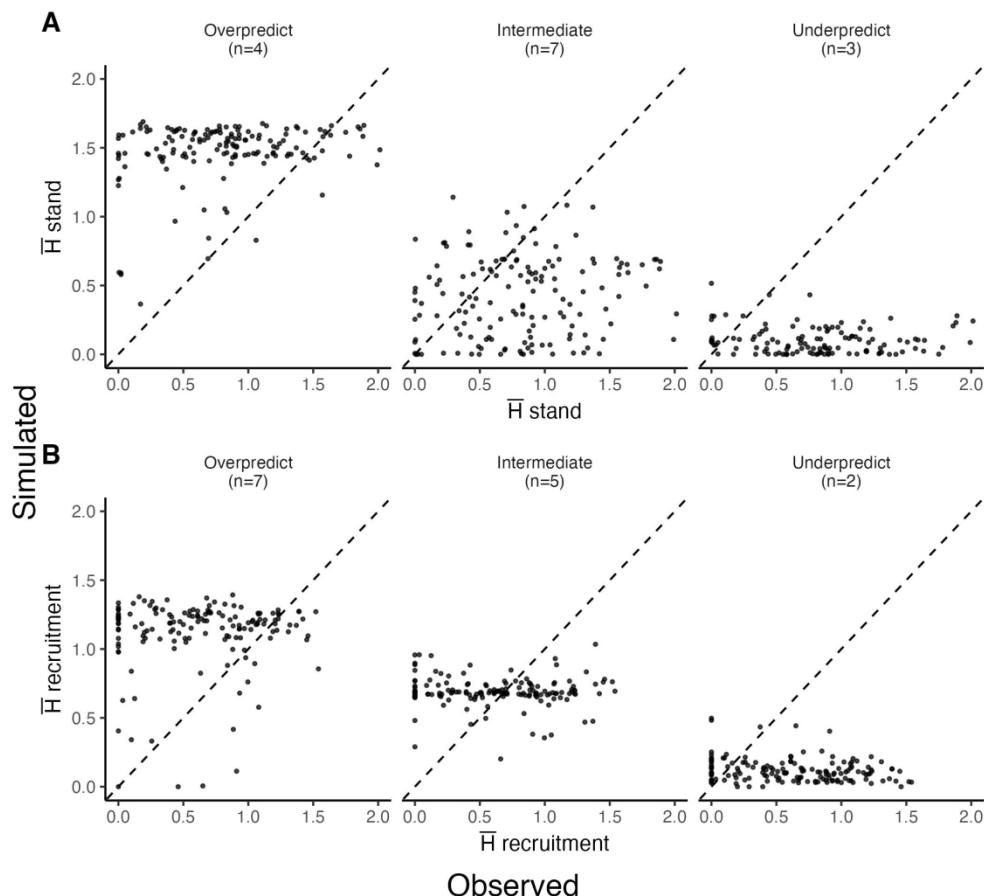


Figure 4. Mean Shannon index across all samples per site for observed and simulated data. Each plot shows one pattern represented by one exemplary model of each category (overpredicted, intermediate and underpredicted). The full data with the grouping of the models are shown in Figures S3 and S4. n indicates the number of models falling in each group. A) Examples for the three trends across models for species diversity at the stand level. B) Examples for the three trends across models for regeneration at the 7 cm threshold (165 sites).

749x674mm (72 x 72 DPI)

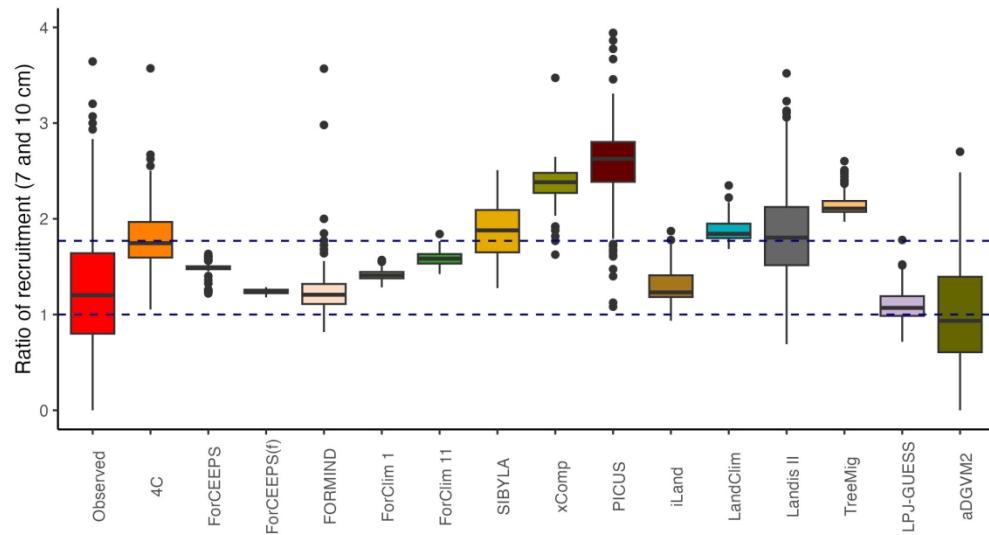


Figure 5. Ratio of tree regeneration rates between the 7 and 10 cm thresholds. Dashed blue lines mark a ratio equal to 1, indicating no decrease in tree regeneration between 7 and 10 cm, and a ratio equal to 1.77, corresponding to the Reineke self-thinning ratio under even-aged conditions.

787x449mm (72 x 72 DPI)

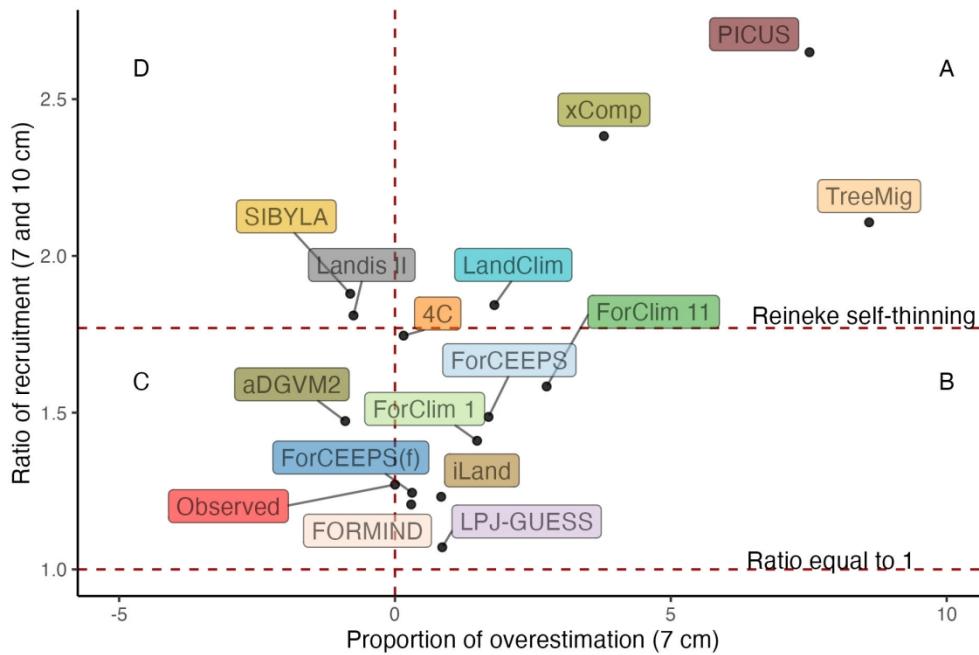
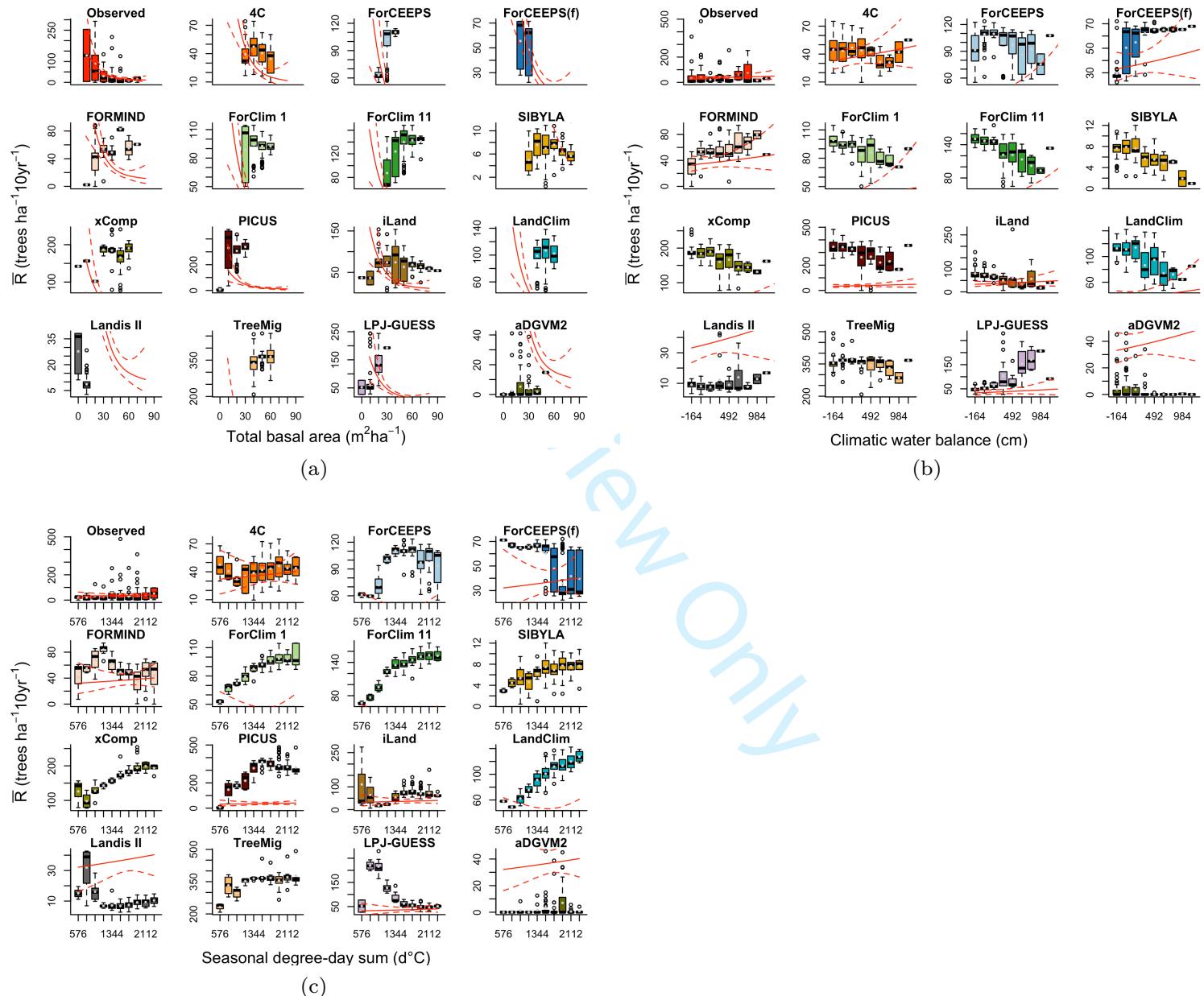


Figure 6. Ratio of tree regeneration between the 7 and 10 cm thresholds (regeneration 7 cm / regeneration 10 cm) and the overestimation proportion at 7 cm ([simulated - observed] / observed) for the mean regeneration per model across sites and samples. The horizontal dashed lines show a ratio equal to 1, indicating no decrease of regeneration between 7 and 10 cm, and a ratio equal to 1.77 corresponding to the Reineke self-thinning line under even-aged conditions. The vertical line indicates zero overestimation.

674x449mm (72 x 72 DPI)



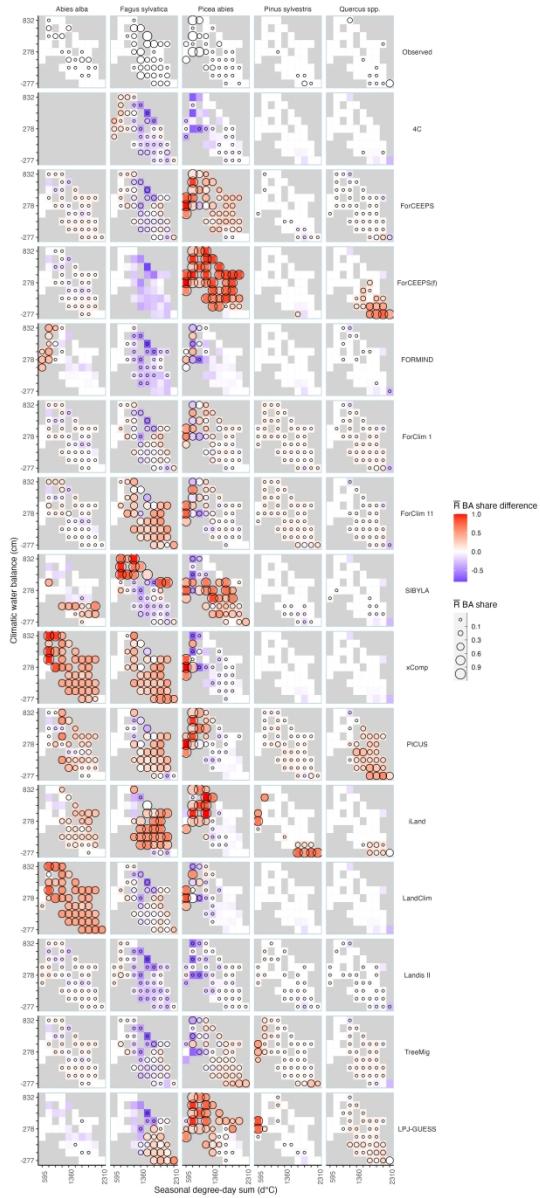


Figure 8. Regeneration levels of the five main species in terms of their basal area share in the observations (top row) and the performance of each model across the environmental gradients (other rows) for the 7 cm regeneration threshold. The values shown are the mean of the 200 samples per site and across the sites in each bin (tile), with ten bins per gradient. The sizes of the circles represent the ratio between the regeneration basal area of the species and the total regeneration basal area for all species. The absence of a circle indicates a zero basal area share in the regeneration, or the absence of regeneration altogether. The color gradient (for the models only) shows the difference between the simulated and observed ratio of regeneration basal area of the species and the total regeneration basal area for all the species.

1012x2249mm (72 x 72 DPI)

Appendix S1

Tree regeneration in models of forest dynamics: a key priority for further research ECOSPHERE

Olalla Díaz-Yáñez, Yannek Käber, Tim Anders, Friedrich Bohn,
Kristin H. Braziunas, Josef Brůna, Rico Fischer,
Samuel M. Fischer, Jessica Hetzer, Thomas Hickler,
Christian Hochauer, Manfred J. Lexer, Heike Lischke,
Paola Mairota, Ján Merganič, Katarina Merganičová, Tobias Mette,
Marco Mina, Xavier Morin, Mats Nieberg, Werner Rammer,
Christopher P.O. Reyer, Simon Scheiter, Daniel Scherrer,
Harald Bugmann

Contents

1 Model reports	2
1.1 4C	2
1.2 ForCEEPS & ForCEEPS(f)	7
1.3 FORMIND	10
1.4 ForClim 1 & ForClim 11	13
1.5 SIBYLA	15
1.6 xComp	21
1.7 PICUS	27
1.8 iLand	30
1.9 LandClim	35
1.10 Landis II	37
1.11 TreeMig	48
1.12 LPJ-GUESS	54
1.13 aDGVM2	60

2 Supplementary figures and tables	67
2.1 Ingrowth levels, tree diversity and mortality in tree establishment	67
2.2 Model traits and model performance	76
2.3 Total ingrowth and individual species regeneration niches	80

1 Model reports

1.1 4C

1.1.1 Authors

Nieberg M., Gutsch M., Reyer C.P.O., Lasch-Born P.

1.1.2 Model

The most complete reference resource for the 4C (v2.2) model and regeneration module is:

- Lasch-Bonn et al. 2020. Description and evaluation of the process-based forest model 4C v2.2 at four European forest sites. DOI: <https://doi.org/10.5194/gmd-13-5311-2020>)

1.1.3 Climate

We used the provided daily ERA5-CHELSA climate data set and complemented the CHELSA time series data regarding relative humidity with the model after [Eccel et al. \(2012\)](#) based on minimum and mean daily temperature to accomodate data needs of 4C. The missing data for January 2nd and 3rd 2013 were filled with data from January 1st and 4th 2013 respectively. For construction

of longer climate time series we looped through the provided time series from 1981 to 2018 up to the maximum simulation length of 2500 years. We amended situations where leap year inconsistencies arised by removing February 29 or duplicating data from February 28. Otherwise, the unaltered daily data were used to run 4C on a daily basis. We applied a constant atmopsheric CO₂ concentration of 380 ppm.

1.1.4 Soil

We used parametrized soils readily available for 4C from the EU soil data base. The link from parametrized soil to soil quality was done by extracting the plant available water storage capacity as the model bucket size of the rooted soil horizons. We excluded bucket size values of lower than 10 cm and larger than 35 cm. The bucket size was reprojected to the range from 0 to 1. In addition we tested the sensitivity of dominant tree height at age 100 years for monospecific stands of all five species simulated to the available soils using model simulations with 4C. The dominant tree height at age 100 years was then also reprojected to the range from 0 to 1 in order to derive a model specific soil quality indicator for all paramterized soils. Then we multiplied this relative model specific soil quality with the relative soil quality derived from the bucket size to the range from 1 to 5. The match between paramterized soil and plot was then done by finding the pair of soil quality with the minimum difference between simuation protocol prescribed soil quality and the soil quality of the paramterized soils as described above.

1.1.5 Topography

Topographical information was not considered in the simulations.

1.1.6 Tree species

We initialized stands with five out of the eleven proposed species/genera: *Fagus sylvatica* (fasy), *Picea abies* (piab), *Betula* spp. (betu; 4C parameters for *Betula pendula*), *Quercus* spp. (quer; 4C parameters for *Quercus robur*) and *Pinus sylvestris* (pisy) since only these are fully parameterized for an application as proposed in the simulation protocol. Each plot was initialized with all five tree species with similar distributions of height classes and number of individuals: 2000 individuals/ha with height around 0.25 m, 1000 individuals/ha with height around 0.75 m and 500 individuals/ha with height around 1.75 m. In total 17500 individuals/ha. *Carpinus betulus*, *Tilia cordata*, *Acer pseudoplatanus*, *Fraxinus excelsior*, *Alnus glutinosa* and *Abies alba* are not fully parameterized in 4C.

1.1.7 Simulations

1.1.7.1 Dispersal and seed input There is no explicitly modelled seed dispersal in a stand and between stands. We applied a constant sapling establishment rate every five years of 2000 individuals/ha for all five species; 10000 individuals/ha in total every five years. The constant sapling establishment provides constant regeneration potential if favorable conditions for sapling growth (especially light and water regime) are present. This approach enables regeneration of other species than only those present in the stand to establish if competition allows. The resulting recruitment patterns are an outcome of the dynamical processes from the sapling stage onward. Seed production and germination were not explicitly simulated.

1.1.7.2 Simulation length and equilibrium criteria To reach equilibrium and obtain dynamics in the equilibrium period, we simulated a maximum of 2500 years. 4C is usually applied for much shorter simulation periods and different technical as well as model-theory related issues arise when simulating longer than 100-150 years. Therefore 2500 years can be seen as the upper limit of continuous simulation in the defined simulation setup. The first 500 years were not sampled for recruitment as we assume the model needs this time to reach equilibrium. In this long simulation period 4C partially projects unrealistically large tree heights resulting in premature termination of the simulation (36 plots) so that not all plots were simulated for the full 2500 years. We started sampling the recruits after the first 500 simulation years from non-overlapping 10 year periods to obtain 200 samples. For shorter simulation periods we produced 200 samples from overlapping 10 year periods. In addition, we excluded those plots in which 4C simulated unrealistically high basal area values of over 150 m² (19 plots). We used this large threshold to give the model maximum flexibility to explore the solution space in these long simulation runs although we acknowledge that this is an unrealistic overestimation that is probably originating from misrepresented density dependent mortality processes in long simulation runs and misrepresented single tree dimensions due to wrong assumptions of tree geometry and biomass share of such old-aged individuals. Furthermore those plots with less than 1000 simulation years (16 plots) before termination were excluded because the sample space after the first 500 years was too small. Additionally we included another subset criterion removing all samples that have a $ba > 90$ to exclude unrealistically high samples without the need to exclude the full site. This resulted in discarding 0.8 % of the samples.

1.1.8 Outputs

We provide 200 samples each for 48 plots and at least 190 samples each for 117 plots while not simulating 35 plots (see reason for excluded plots in “Simulation length and equilibrium criteria”). These stem from 10 year periods that overlap partially (for < 2500 simulation years) or do not overlap (for 2500 simulation years). The number of recruits (r.trees) is the sum of individuals that cross the threshold diameter (7/10 cm) in the 10 year sample period per hectare that have not already died until the sample year. The recruit’s basal area (r.ba) is the sum of the basal area per hectare of the trees that cross the diameter threshold that have not already died until the sample year. The total stand basal area (ba) is the basal area per hectare of all trees in the stand > 0 cm diameter at breast height (no lower boundary) at the sample time per species. We provide the data for both diameter thresholds, 7 and 10 cm. In over 90 % of the decades no recruitment takes place.

4C (v2.2) is usually not used for simulations of forest equilibrium dynamics, potential natural vegetation assessment and long simulation periods but rather shorter-term simulations (up to 100-150 years). In addition, only five species out of the pool of eleven potentially relevant species here are parameterized in 4C so that projections of the potential natural vegetation done with 4C need to be considered in the light of model uncertainty. For example, the spatial homogeneity in the model leads to cyclic recruitment patterns. When the adult stands dies off, young individuals are recruited into the larger diameter classes because more light reaches lower canopy layer since larger (adult) trees do not anymore prevent the light reaching lower canopy layers. This results in increased growth of the regeneration layer during those times. In most stands *Fagus sylvatica* dominates the stand after some time but occasionally individuals of other species, like *Picea abies* or *Quercus* spp. cross the recruitment diameter threshold.

1.2 ForCEEPS & ForCEEPS(f)

1.2.1 Author

Xavier Morin & François de Coligny

1.2.2 Model

The most complete reference resource for the ForCEEPS & ForCEEPS(f) model and regeneration module are:

- Morin X., de Coligny F., Bugmann H., Limousin J.-M., Ourcival J.-M., Martin-StPaul N., Simioni G., Cailleret M., Prevosto B., Toïgo M., Vennetier M., Cateau E., Guillemot J. 2021. Beyond forest succession: a gap model to study ecosystem functioning and tree community composition under climate change. *Functional Ecology*. 35, 955-975.
- Jourdan M., Dreyfus, P., Riond C., Cordonnier T., Cornet B., de Coligny F., Morin X. 2021. Managing mixed stands can mitigate severe climate change impacts on French alpine forests. *Regional Environmental Change*. 21, 78.
- Morin X., Damestoy T., Toïgo M., Jactel H., Castagnéryrol B. Meredieu C. 2020. Using forest gap models and experimental data to explore long-term effects of tree diversity on the productivity of mixed planted forests. *Annals of Forest Science*. 77 <https://doi.org/10.1007/s13595-020-00954-0>.

1.2.3 Climate

ForCEEPS requires monthly data. A 4000-yr long time-series was obtained for each site by randomizing available years from the CHELSA database, with mean monthly temperature and monthly sum of precipitations (see 1.2.2).

1.2.4 Soil

Soil quality is represented by a continuous value between 1 and 5, as provided (see 1.2.3) Soil Field Capacity (SFC) was assessed as follows: $SFC = 8 + (X-1)*(25-8)/(5-1)$, in which X is the ‘soil_qual’ value in the provided ‘soil_quality_200.csv’ file.

1.2.5 Topography

Not considered.

1.2.6 Simulations

In each site, the simulation consists of 4000 yr-long simulation of 1ha of forest (ie. 10 patches of 1000 m²), starting from bare-ground. There was no dispersal between patches. Two kinds of simulations have been carried out in each of the 200 sites. In a first set of simulations, regeneration dynamics was simulated with a seed rain from a potential species list, thus independent from the actual forest composition, as classically done in gap models. Under this design, a number of seedling is randomly chosen for each patch, and a species identity is randomly sorted from a potential species list (defined at the start of simulation) assigned to each seedling. Then each seedling

tries to colonize the site depending on whether the suitability between species requirements and site conditions (climate, light, soil quality, browsing index). In a second set of simulations, regeneration dynamics was simulated with a seed rain whose composition depended on the actual forest composition. Under this design, a number of seedling is randomly chosen for each patch, but the species identity is assigned according to species relative abundance (calculated across all 10 patches according to species biomass). For instance, if at year i the adult trees of species j represent 60% of the total biomass across the 10 patches, then 60% of the seedlings that try to colonize the patches at year i will belong to species j. Then, similarly to the first design, each seedling tries to colonize the site depending on whether the suitability between species requirements and site conditions. In the two sets of simulations, simulations have been run for both thresholds for regeneration (7 and 10 cm).

1.2.7 Outputs

Outputs consisted in species specific regeneration rates per decade and per ha, by considering the last 2000 years of simulations to ensure that pseudo-equilibrium state is reached. Results were aggregated for every 10 years for the last 2000 years at the 1ha-level. Variables: - Site: numbers corresponding to the climate data file - Sample: one per decade, for a total of 2000 years = identified from 1 to 200 - Species: name specified as requested (eg. ‘fasy’) - r.trees: number of trees outcrossing the threshold (7 or 10 cm) during the decade, for the 10 patches simulated (= 1ha), per species - r.ba: summed basal area of the r.trees, per species - ba: summed basal area of all trees per species at the end of the decade - dbh: threshold used

1.3 FORMIND

1.3.1 Authors

Samuel M. Fischer, Friedrich Bohn, Rico Fischer

1.3.2 Model

A full description of the model can be found in Fischer et al (2016) . The basis for the parameterization we applied is described in Bohn et al (2014):

- Bohn, F. J., Frank, K., & Huth, A. (2014). Of climate and its resulting tree growth: Simulating the productivity of temperate forests. *Ecological Modelling*, 278, 9–17. <https://doi.org/10.1016/j.ecolmodel>.
- Fischer, R., Bohn, F., Dantas de Paula, M., Dislich, C., Groeneveld, J., Gutiérrez, A. G., Kazmierczak, M., Knapp, N., Lehmann, S., Paulick, S., Pütz, S., Rödig, E., Taubert, F., Köhler, P., & Huth, A. (2016). Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests. *Ecological Modelling*, 326, 124–133. <https://doi.org/10.1016/j.ecolmodel.2015>.

1.3.3 Climate

To incorporate climate, we used the data set ‘daily_era5_chelsa_200.csv’, which contains climate data on a daily resolution. We considered the covariates precipitation, temperature, and irradiance, the latter of which we converted to units used in Formind. Adding on to these data, we used a

constant and site-independent potential evapotranspiration (PET) value of $4\frac{\text{mm}}{\text{d}}$ and assumed that the mean day length was 12h at all sites.

Since we simulated longer time periods than the 38 years for which we had climate data, we used the provided data in a randomized fashion. For each simulation year, we applied the climate data of a randomly selected year. We neglected the existence of leap years, which led to a minor time shift in climate data of later years.

1.3.4 Soil

We modelled soil conditions by using the provided soil quality values as proxy for soil depth. To that end, we scaled the provided values to the typical range of soil depth values used in Formind for temperate forests. Specifically, we multiplied the soil quality values by factor 0.5 to obtain the modelled soil depth in meters.

1.3.5 Topography

We did not use any topography data.

1.3.6 Tree species

We considered each of the 11 species in the protocol explicitly. We chose the model parameters for the individual species based on the work by Bohn et al (2014). For species that were not covered by Bohn et al (2014), we used parameter values of similar species with known parameters. The species composition was not predefined and evolved solely from the seed input and the model dynamics.

1.3.7 Simulations

1.3.7.1 Simulation area We simulated a forest area of 1ha, which was divided into 25 patches of size 20m × 20m. No interactions between trees from different patches were considered in the model. That is, the 25 patches were independent replicates of one another.

1.3.7.2 Seed generation and dispersal We modelled seed generation as global external seed rain. Specifically, we assumed that for each species, there is a constant seed inflow of $250 \frac{\text{seeds}}{\text{ha}\cdot\text{yr}}$, which is distributed evenly among the 25 individual patches. We supposed that all seeds promptly establish to seedlings with initial stem diameters at breast height (DBH) of 5cm.

1.3.7.3 Simulation length and sampling approach We started the simulation at each site with a burn-in phase of 1000 years and collected samples of the output variables in subsequent years. The length of the burn-in phase was chosen based on earlier experiences with the model. To confirm that the model reached its limiting behaviour, we plotted the evolution of the output variables (e.g. the basal area) at selected sites over time and verified that no trends were visible after the burn-in phase.

We computed the output variables in equidistant time intervals of 60 years. Before taking a sample, we saved parts of the model's state. Then, we simulated the forest for 10 years and computed the output variables, comparing the model's current state with the state saved earlier (see below). After taking the sample, we simulated the forest for another 50 years to reduce correlations between samples. Then, we repeated the procedure until the desired number of 200 samples was collected.

1.3.8 Outputs

To determine the number and basal area of recruits, we compared the set of large trees before and after the regeneration time period of 10 years. Prior to simulating the forest for the regeneration time period, we saved the internal IDs of all trees with DBHs above the thresholds of 0.07m or 0.1m, respectively. We stored these IDs in set data structures $S_{0.07}^{\text{prior}}$ and $S_{0.1}^{\text{prior}}$, corresponding to one threshold value each. After the regeneration time, we repeated the procedure, obtaining tree ID sets $S_{0.07}^{\text{posterior}}$ and $S_{0.1}^{\text{posterior}}$. We then computed the set differences $S_{0.07}^{\text{recruits}} = S_{0.07}^{\text{posterior}} \setminus S_{0.07}^{\text{prior}}$ and $S_{0.1}^{\text{recruits}} = S_{0.1}^{\text{posterior}} \setminus S_{0.1}^{\text{prior}}$ and determined both the number and cumulative basal area of the trees in $S_{0.07}^{\text{recruits}}$ and $S_{0.1}^{\text{recruits}}$ by considered species, respectively.

Finally, we computed the total basal area for each considered species by adding up the basal areas of all corresponding individuals. Note that the initial DBH of seedlings in our model was 0.05m. Consequently, smaller trees were not included in the basal area computation.

1.4 ForClim 1 & ForClim 11

1.4.1 Authors

Yannek Käber, Harald Bugmann

1.4.2 Model

The most complete reference resource for the ForClim 1 & ForClim 11 model and regeneration module is:

- Bugmann, H. (1994). On the ecology of mountainous forests in a changing climate: A simulation study.
- Huber, N., Bugmann, H., & Lafond, V. (2020). Capturing ecological processes in dynamic forest models: Why there is no silver bullet to cope with complexity. *Ecosphere*, 11(5).
<https://doi.org/10.1002/ecs2.3109>

1.4.3 Climate

Temperature and precipitation were aggregated to monthly means and sums, respectively, including their standard deviation. These values served as input for the weather generator within the model. The weather generator simulates wheather conditions in an annual time step based on the monthly average temperatures and precipitations sums.

1.4.4 Soil

Soil data were translated to bucket sizes values based on the assumption that the lowest soil quality value of 1 refers to a bucket size of 10 cm and the highest soil quality of 5 refers to a bucket size of 35 cm. Specifically we used the following formula to calculate the bucket size $\text{bucket_size} = 10 + (\text{soil_quality}-1) / 4 * 25$.

1.4.5 Topography

Topographic information was included by specifying the kSIAsp parameter. This parameter defines a value between -2 and 2 which affects the calculated Potential Evapotranspiration within the model. For details on the calculation of this value see Käber et. al. 2021.

1.4.6 Tree species

All species required were simulated: *Fagus sylvatica*, *Picea abies*, *Abies alba*, *Carpinus betulus*, *Tilia cordata*, *Acer pseudoplatanus*, *Betula spp.*, *Fraxinus excelsior*, *Quercus spp.*, *Alnus glutinosa*, *Pinus sylvestris*.

In addition the following species were simulated: *Acer campestre*, *Acer platanoides*, *Alnus incana*, *Alnus viridis*, *Castanea sativa*, *Corylus avellana*, *Larix decidua*, *Pinus montana*, *Populus nigra*, *Populus tremula*, *Salix alba*, *Sorbus aria*, *Sorbus aucuparia*, *Taxus baccata*, *Tilia platyphyllos*, *Ulmus glabra*, *Pinus cembra*

1.4.7 Simulations

All simulations were run according to the protocol. For each site we simulated on 16 ha (i.e., 100 patches of 0.08 ha). We defined a spin up phase of 1000 years. After the spin up we simulated 200 years of forest dynamics.

1.4.8 Outputs

We sampled in 10 year intervals between the simulation years 1010 to 1200, which resulted in 16 times 20 intervals each with one ha.

1.5 SIBYLA

1.5.1 Authors

Ján Merganič, Katarína Merganičová, Marek Fabrika, Peter Valent

1.5.2 Model

SIBYLA is the simulator of forest biodynamics. It belongs to the category of semi-empirical tree growth simulators. It consists of the set of mathematical models and algorithms that are transformed into an integrated software package SIBYLA Suite. The model has been developed at the Department of Forest Management and Geodesy, Technical University in Zvolen, Slovakia. The basis of the model was the modelling principle and algorithms implemented in SILVA 2.2 (Pretzsch 1992, Kahn 1994). The model is sensitive to climatic factors (length of growing season, mean temperature during growing season, annual temperature amplitude, and total precipitation during growing season). The climatic factors modify height and diameter growth potential, and consequently tree increment. The climatic factors also influence tree regeneration model.

The most complete reference resource for the SIBYLA model and regeneration module are:

- Fabrika (2005) “Simulátor biodynamiky lesa SIBYLA.” *Koncepcia, konštrukcia a programové riešenie. Habilitačná práca, Technická univerzita vo Zvolene*
- Materials available in <http://etools.tuzvo.sk/sibyla/english/model.htm>.

1.5.3 Climate

We used CHELSEA monthly and daily time series from 1981 to 2018 to derive climate variables necessary for simulations with SIBYLA as follows: Monthly average temperature (tas) values were used to derive the SIBYLA site-specific characteristic called the „annual temperature amplitude” (labelled as TAMPL or s5 in the model) that was calculated as a difference between the maximum and minimum monthly average temperatures. Monthly average temperature (tas) values for months

April to September were used to derive the SIBYLA site-specific characteristic called the „average temperature during the growing season” (labelled as TEMP or s6 in the model). Monthly precipitation sums (pr) for months April to September were summed up to derive the SIBYLA site-specific characteristic called the „precipitation total during the growing season” (labelled as PRECIP or s8 in the model). Daily average temperature (tas) values were used to derive the SIBYLA site-specific characteristic called the „length of the growing season” (labelled as DAYS or s4 in the model).

1.5.4 Soil

The information on site-specific soil quality (file name: soil_quality_200.csv) was used to define two SIBYLA site-specific characteristic called the „soil moisture” (labelled as MOIST or s7 in the model) and the „soil nutrient supply” (labelled as NUTR or s3). Since these two characteristics can obtain values in the range from 0 to 1, the provided values on soil quality were converted to this range by dividing each value with 5 (the maximum). We used the long-term average of CO₂ concentrations in air that represented the period 1981 – 2018. The long-term average of CO₂ was calculated from the annual CO₂ data obtained from Mauna Loa observations (Keeling et al. 1976). Similarly, we used the long-term NO_x concentration representing the time period 1981-2018. This was obtained by averaging annual values of NO_x concentration in air calculated using the equation by Kahn (1994), where the year is the driving variable.

1.5.5 Topography

The growth simulator SIBYLA uses the model of ecological classification applied in the growth simulator SILVA 2.2, which was derived by Kahn (1994). Site is specified using ecological site

characteristics describing climate, air, and soil, also called site variables:

- s1 (N2O) ... NOx concentration in air (ppb)
- s2 (CO2) ... CO2 concentration in air (ppm)
- s3 (NUTR) ... soil nutrient supply (relative value in the range from 0 to 1)
- s4 (DAYS) ... number of days in the vegetation period (days with daily mean temperature above 10°C)
- s5 (TAMPL) ... annual temperature amplitude (the difference between minimum and maximum monthly temperature in °C)
- s6 (TEMP) ... daily mean temperature in the vegetation period in °C (from April to September)
- s7 (MOIST) ...soil moisture (relative value in the range from 0 to 1)
- s8 (PRECIP) ... precipitation amount in the vegetation period in mm (from April to September)
- s9 (ARID)... aridity index according to de Martone in mm.°C-1 They directly influence the production capacity of a stand (tree height and diameter increment).

1.5.6 Tree species

At the beginning of the simulations we determined balanced tree species composition of all 12 tree species pre-defined in the „Regeneration workshop protocol”. It means that at the beginning each species was represented by the same volume, mean tree diameter, mean tree height, and age.

1.5.7 Simulations

The simulations were performed following the “Regeneration workshop protocol”. Hence, we simulated the pre-defined 200 sites with the model SIBYLA under current climate. The simulations were performed with the help of two software tools for automated handling of multiple sites to speed up the process called DBCreator and Automat (Valent et al. 2018).

1.5.7.1 Simulation area and site We set the simulation area to 1ha, i.e. each site was represented by an area of 1 ha, to enable a sufficiently large area for species interactions. Larger areas were not possible due to the restrictions in the number of trees per plot (the maximum number of trees per plot is 6,000). Due to the restriction in the maximum number of trees per plot we established a 40-year-old forest stand at each site with the balanced tree species composition as defined above. We used site-specific elevation and aspect obtained from the information on topography and elevation (file name: topo_data_200.csv).

1.5.7.2 Simulation length Each site was simulated for 3,000 years. Based on the analysis of the temporal development of multiple stand variables including number of trees, mean tree and stand characteristics, the simulation time until the equilibrium was set to 500 years. Hence, the first 500 simulated years were excluded from the results. The submitted output includes the required information per decade during the 2,500 years.

1.5.7.3 Dispersal Relevant aspects of the dispersal (if any) and seed input parameters in your model The regeneration model operates if there are mature trees able to produce seeds. Hence, the initial forest stand had to be established to allow the reproduction and regeneration module. The

seed yield of individual trees is predicted in relation to species-specific characteristics including tree age interval of seed production and fertility (age of start, optimum and end), seed production per adult tree, transmission distance, seed purity, seed quality, germination capacity, absolute weight of seeds. Seed tree production is further modified with the regulators and reduction factors, which describe other conditions necessary for the germination, survival, and the establishment of the next generation, e.g. stand canopy closure (Fabrika 2005, Merganič and Fabrika 2009, Fabrika et. al. 2009, 2011).

1.5.8 Outputs

To derive the required information, we worked with the model output for individual trees per decade. First, we calculated basal area of individual trees from the simulated diameter at breast height using the equation for the area of a circle. Total basal area (ba) was obtained by summing tree basal areas per species and decade. The number of recruits (r.trees) per species was derived by counting the trees that exceeded the diameter threshold (7 or 10 cm) in the specific time. The basal area of recruits (r.ba) per species was calculated by summing tree basal areas of recruits per species. NA was assigned if no trees of a specific species occurred in the period.

1.5.9 References

Fabrika M., Merganič J., Merganičová K., 2009: Natural regeneration density model developed for the purposes of the individual-tree growth simulator. Acta Facultatis Forestalis Zvolen, 51(3), p. 125-137, ISSN: 0231-5785

Fabrika, M., (2005): Simulátor biodynamiky lesa SIBYLA, koncepcia, konštrukcia a programové

riešenie. Habilitačná práca. Technická univerzita vo Zvolene, 238 p.

Kahn, M., (1994): Modellierung der Höhenentwicklung ausgewählter Baumarten in Abhängigkeit vom Standort. Forstliche Forschungsber. München, Vol. 141, 221 p.

Keeling C. D., R. B. Bacastow, A. E. Bainbridge, C. A. Ekdahl Jr., P. R. Guenther, L. S. Waterman, J. F. S. Chin 1976: Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii. Tellus. XXVIII, 6: 538-551, <https://doi.org/10.1111/j.2153-3490.1976.tb00701.x>

Merganič J., Fabrika M., 2009: Determination of climatic-site strata based on rectified climatic rasters for the purposes of natural regeneration density model. Acta Facultatis Forestalis Zvolen, 51(3), p. 113-124, ISSN: 0231-5785

Merganič, J., Fabrika, M., Merganičová, K., 2011: Submodel of height-diameter function for climatic-site strata of the model of natural regeneration in the growth simulator SIBYLA. Acta Facultatis Forestalis Zvolen, 53(1), p.155-168, ISSN: 0231-5785

Pretzsch, H., (1992): Konzeption und Konstruktion von Wuchsmodellen für Rein- und Mischtbestände. Forstliche Forschungsberichte München, Nr.115, 358 p.

Valent, P., Merganič J., Výbošťok J. 2018: Optimus – Software for Multi-criteria analysis. TU Zvolen.

1.6 xComp

1.6.1 Authors

Tobias Mette, LWF.

1.6.2 Model

The forest growth model xcomp was developed to simulate gap dynamics of monospecific stands with a strong emphasis on position-dependent individual tree competition. The model consists of three units: (1) the main growth phase which determines the growth and mortality of each individual tree above 5 cm dbh, (2) the old-growth phase where individual trees exceeding a certain dbh-limit decline in vitality and eventually die, and (3) the regeneration phase where seedling growth and mortality is modelled depending on light availability. The primary focus of the model was to establish a set of robust algorithms that can be flexibly calibrated for different species attributes. The model was used to simulate single-species stand development cycles for *Nothofagus pumilio* in Patagonia (Mette 2014), for European beech in NE-Bavaria (Mette et al. 2016) and a Norway spruce stand with different mortality algorithms (Bugmann et al. 2019). For the current article the model was (4) connected to a climate sensitive species distribution model to modify growth and mortality in dependence of the site conditions, and (5) extended from single to multi-species simulations (c.f. Section 1.6.7).

The most complete reference resource for the xComp model and regeneration module are:

- Mette, T. (2014). xcomp Tutorial (v2014_08). Technical Report, TUM. DOI: 10.13140/RG.2.2.32524.59529
- Mette, T. (2014) Using Reineke to self-calibrate mortality in individual tree based forest growth modelling. Beiträge zur Forstwissenschaftlichen Tagung, Dresden, 17.-20.09.2014. 10.13140/RG.2.2.11553.07527

1.6.3 Climate

The species parameterisation in xcomp is sensitive to long-term climate, i.e. 30-year averages of summer temperature, summer precipitation and winter temperature. Reactions to daily, monthly or periodic extremes are not implemented as deterministic but as stochastically processes functionally dependent on the long-term climate. The climate sensitivity of growth, mortality and regeneration is described in Section 1.6.7.

1.6.4 Soil

Soil moisture class MOIST was used to modify the climate sensitivity via summer precipitation PJJA by $PJJA(\text{mod}) = PJJA + 15 * \text{MOIST} - 45$ (in mm).

1.6.5 Topography

Topography is no input parameter in xcomp. It may only effect the simulation results indirectly through climate or soil parameters.

1.6.6 Tree species

From the suggested species spectrum *Alnus glutinosa* and *Fraxinus excelsior* were not included in the species mix. *Quercus spec* was split into *Q. petraea* and *Q. pubescens*, but united in the validation to meet the protocol.

1.6.7 Simulations

The simulation was carried out according to the protocol by Bugmann et al. (2022). I.e. the simulation is initiated for mixed species stands of 1 ha size at 200 sites of with given climate and soil conditions. The initial species mix consists of up to 10 species of equal shares with an initial root mean square diameter dq of 8 cm. Xcomp simulates forest growth in 5 year intervals. Data are validated after having reached a climate dependent equilibrium – which is typically the case after 750 years (= 150 simulation periods). Starting with year 760, data are read out for 200 10 year intervals (until year 2750, simulation period 550). Further simulation relevant model details:

- (1) The main growth phase is determined by the processes growth and mortality. Mortality is determined by the Reineke exponent rke and the stand density sdi and calibrated over a range of 8 cm to 50 cm dq (= root mean square diameter) so that stem number N equals: $N = sdi * (dq/25)^{rke}$. A tree's resource acquisition area res (m^2) is approximated as $res = 1e4 / (sdi * (dbh/25)^{rke})$. As the Reineke exponent is fixed to $\square 1.6$, species with higher/ lower sdi have a lower/ higher resource acquisition area for a given dbh. Potential dbh growth dd is calculated by multiplying a reference dbh growth with a site specific temperature dependent modifier. The individual dbh growth in the stand is further modified by the neighbourhood competition.
- (2) The old growth phase sets in when a tree passes a certain old growth dbh (ogDbh) which depends on the site conditions. Having passed this threshold dbh a tree is assigned a mortality period that assumes a value not greater than a site independent maximum ogPer. The maximum old growth time span ranges from 100 years (birch) to 350 years (fir). During the old growth phase dbh growth declines and more light becomes available for regeneration.
- (3) The regeneration phase allows seedling development on a 1x1 m grid over the entire stand.

First, a light availability matrix (1x1 m) is calculated from a cumulative crown shadow which decreases with increasing distance to the tree center. Next, regeneration matrices (1x1 m) are created for each of the stand's species. In the initial period, the regeneration matrix of each species is empty, but at the end of each period, a light dependent dbh growth and mortality is calculated for each 1m raster cell. The dbh growth is scaled so that under full light 5 cm dbh is reached in a defined regeneration period (regnPer) that depends on the site conditions. The default value is set to 15 years and equal for all species. Dbh growth is modified by the light availability and the species' regeneration light demand. This is a decisive difference between the species and favours shadow-tolerant trees like fir and beech in the absence of large gaps. However, each raster cell is afflicted with a species bonus or malus that provides local variation to the global regeneration parameters. Mortality in the regeneration raster occurs as a random process, the likelihood of which depends on the climate.

(4) The climate sensitivity was calculated by methods of species distribution modelling. Effectively species abundance in European NFI-plots (Mauri et al. 2017) was related to climate of the plots (Karger et al. 2017) using generalized additive models as in Thurm et al. (2018). Input variables were summer temperature mean (Jun-Aug), winter temperature mean (Dec-Feb) and summer precipitation sum (Jun-Aug). Abundance was transformed into prevalence () so that the values for each species can assume a range from 0 to 1. The prevalence values modify dbh growth dd, old growth dbh ogDbh, the regeneration period regnPer, and the regeneration mortality regnMort.

(5) The extension from single to multiple species simulation introduced no technical changes to the single-species growth and mortality behaviour. Competition as the main interaction between trees depends on the crown size and density of the neighbouring trees and can turn out an dis/advantage for one or the other species. But species calibration parameters were not changed due to mixture-

specific interaction.

1.6.8 Outputs

Since the simulation output tables contain all biometric variables on an individual tree basis, all required outputs can be derived directly from the output tables. Being a very deterministic model what typically happens in xcomp is that the equilibria consist of one or two species. Shade tolerance favours fir and beech in the absence of large gaps creating disturbances. The second most important parameter is the site suitability which is estimated by the modelled species prevalence and affects growth and thereby competitiveness. As it is the first time that xcomp is used in a multiple species application there is room for improvements. The automation of a robust calibration of the main growth phase is a precondition to investigate how for instance larger disturbances, a weighted regeneration potential oriented at the main stand, higher or selective mortality of the regeneration affect the species composition and turnover.

1.6.9 References

1. Bugmann, H., Seidl, R., Hartig, F., Bohn, F., Brůna, J., Cailleret, M., ... & Reyer, C. P. (2019). Tree mortality submodels drive simulated long-term forest dynamics: Assessing 15 models from the stand to global scale. *Ecosphere*, 10(2), e02616.
2. Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... & Kessler, M. (2017). Climatologies at high resolution for the earth's land surface areas. *Scientific data*, 4(1), 1-20.
3. Mauri, A., Strona, G., & San-Miguel-Ayanz, J. (2017). EU-Forest, a high-resolution tree

- occurrence dataset for Europe. *Scientific data*, 4(1), 1-8.
4. Mette, T., Falk, W., Blaschke, M., Förster, B., Walentowski, H. (2016). Modelling natural forest dynamics in individual tree based forest growth simulators. *Beiträge zur Forstwissenschaftlichen Tagung*, Freiburg, 26.-29.09.2016. 10.13140/RG.2.2.29169.15201
 5. Thurm, E. A., Hernandez, L., Baltensweiler, A., Ayan, S., Rasztovits, E., Bielak, K., ... & Falk, W. (2018). Alternative tree species under climate warming in managed European forests. *Forest Ecology and Management*, 430, 485-497.

1.7 PICUS

1.7.1 Authors

Christian Hochauer, Manfred J. Lexer

1.7.2 Model

The core elements of PICUS are described in Lexer and Hönniger (2001) and Seidl et al. (2005).

Specific details regarding the regeneration module can be found in Woltjer et al. (2008).

The most complete reference resource for the SIBYLA model and regeneration module are:

- Lexer MJ, Hönniger K (2001) A modified 3D-patch model for spatially explicit simulation of vegetation composition in heterogeneous landscapes. *For Ecol Manag* 144:43–65. doi: [10.1016/S0378-1127\(00\)00386-8](https://doi.org/10.1016/S0378-1127(00)00386-8)
- Seidl R, Lexer MJ, Jäger D, Hönniger K (2005) Evaluating the accuracy and generality of a hybrid patch model. *Tree Physiol* 25:939–951. doi:[10.1093/treephys/25.7.939](https://doi.org/10.1093/treephys/25.7.939)

- Woltjer M, Rammer W, Brauner M, Seidl R, Mohren GMJ, Lexer MJ (2008) Coupling a 3D patch model and a rockfall module to assess rockfall protection in mountain forests. *J Environ Manag* 87:373–388. doi:10.1016/j.jenvman.2007.01.031

1.7.3 Climate

We used the provided monthly climate data (monthly_era5_chelsa_200.csv). The provided solar radiation rsds was converted to MJ per m². Additionally, vpd was calculated.

1.7.4 Soil

Based on the provided continuous value for soil quality between 1 and 5 minimum (soil quality = 1) and maximum (soil quality = 5) values for soil water holding capacity (80mm, 200mm), plant available Nitrogen (40 kg ha⁻¹ yr⁻¹, 100 kg ha⁻¹ yr⁻¹) and soil pH (4.2, 5.6) were defined. Linear interpolation was used to scale between the minimum and maximum values.

1.7.5 Topography

Not considered.

1.7.6 Tree species

The species composition was defined by the given species from the protocol: *Fagus sylvatica*, *Picea abies*, *Abies alba*, *Carpinus betulus*, *Tilia cordata*, *Acer pseudoplatanus*, *Betula spp.*, *Fraxinus excelsior*, *Quercus spp.*, *Alnus glutinosa* and *Pinus sylvestris*. For *Betula spp.* *Betula pendula* was simulated, for *Quercus spp.* the species *Quercus petraea* and *Quercus robur* where used. The

establishment from bare ground was achieved via external seed input (see also in section seed dispersal), so depending on the environment and random components all the given species had a chance to grow into the respective simulation. In addition, on site production of seed is considered depending on age and size of potential parent trees.

1.7.7 Simulations

1.7.7.1 Simulation Area In PICUS the area is defined via the number of 10 m by 10 m patches. For this simulation we used an area of 12 by 12 patches resulting in an area of 1.44 ha. This size of the simulated forest provides relatively stable species composition & structure and reasonable simulation time.

1.7.7.2 Seed dispersal For the experiment seed input to the simulated forest happens via 2 pathways: (1) external seed input to every patch in any simulated year (fixed number of seed per patch per species). (2) locally produced seed by parent trees growing in the simulated forest. Seed production depends on species, and tree size and a random component accounting for mast years. Seed dispersal depends on whether a species is anemochorous or zoolochorous. If anemochorous tree height and seed characteristics determine the maximum dispersal range and the shape of the dispersal function. Based on germination rate, the environmental factors (temperature, soil moisture, Nitrogen, light) and the amount of seeds in combination with a random component seedling establishment is calculated per patch in any given year. Seedling cohorts grow through 4 height classes and are then eventually recruited into the tree population.

1.7.7.3 Simulation length We decided to use the temporal sampling. We first run the model into a quasi equilibrium and continue afterwards with another 2000 years. The simulation time until equilibrium was chosen with 600 years from experience in other projects and a few simulations with test sites where 600 years was a sufficiently long simulation time to reach an equilibrium state.

1.7.8 Output

Output consisted of species-specific regeneration rates per decade and per ha. Oak species were aggregated. A sample was taken for every 10 years for a total period of 2000 years once the model had reached equilibrium (see 1.7.7.3). Ingrowth was determined by counting the trees passing the 7cm and 10cm DBH thresholds every 10 years. Basal area of the ingrowth was determined at the end of each decade.

1.8 iLand

1.8.1 Authors

Werner Rammer, Kristin Braziunas

1.8.2 Model

The most complete reference resource for the iLand regeneration module is:

- Seidl, R., Spies, T. a. T. A. T. a., Rammer, W., Steel, E. A. A., Pabst, R. J. R. J. R. J., & Olsen, K. (2012). Multi-scale Drivers of Spatial Variation in Old-Growth Forest Carbon

Density Disentangled with Lidar and an Individual-Based Landscape Model. *Ecosystems*, 15(8), 1321-1335. <https://doi.org/10.1007/s10021-012-9587-2>

1.8.3 Climate

Climate inputs for iLand are at daily temporal resolution. We used daily climate from the CHELSA dataset for maximum and minimum temperature, precipitation, and surface shortwave radiation. We calculated daily vapor pressure deficit (VPD) from these inputs assuming daily minimum temperature is dewpoint temperature, a reasonable assumption except under very arid conditions (Campbell and Norman 2000). We calculated vapor pressure from temperature using the Tetens formula with different coefficients for temperatures above or below 0° C (Murray 1967). We then followed methods outlined by the Food and Agriculture Organization of the United Nations (Allen et al. 1998) to calculate vapor pressure deficit as the difference between average daily saturation vapor pressure (derived from minimum and maximum temperature) and ambient vapor pressure derived from dewpoint temperature. Some daily VPD values were negative because there were days in the CHELSA dataset where minimum temperature exceeded maximum temperature; negative VPD values were set to 0.

As recommended by workshop organizers, we filled in daily values for two missing days in 2013 (Jan 2 and Jan 3) with daily values for Jan 1 and 4, respectively.

Atmospheric CO₂ concentration was set to 380 ppm and latitude was set to 50° N for all sites. Annual climate year was drawn randomly with replacement from the full dataset (1981-2018) for each site for the duration of the simulation.

1.8.4 Soil

Inputs for soil depth, texture, and fertility (plant available nitrogen) were extracted from a previously compiled Europe-wide dataset (unpublished dataset). We averaged values in a 1 km buffer around blurred site coordinates provided by workshop organizers. Relationships between ranked site quality provided by workshop organizers and soils data from our Europe-wide dataset were weak (Pearson's $|r|$ ranged from 0.04 to 0.34), but followed expected trends with lower sand content, higher water holding capacity, and higher fertility associated with higher site quality. Albeit the weak correlation, we chose to use our available data set because we could avoid additional assumptions as it included all site variables needed for the model.

1.8.5 Topography

Topographic information was not used.

1.8.6 Tree species

Tree composition is defined at the species level. All individually requested species (*Fagus sylvatica*, *Picea abies*, *Abies alba*, *Carpinus betulus*, *Tilia cordata*, *Acer pseudoplatanus*, *Fraxinus excelsior*, *Alnus glutinosa*, *Pinus sylvestris*) were included. For the two taxa where species was not defined, we identified a representative species (*Betula pendula* for *Betula* spp. and *Quercus robur* for *Quercus* spp.). No other species were included in model simulations.

1.8.7 Simulations

1.8.7.1 Simulation area The simulation extent for each site was 1000 m x 1000 m (total area = 100 ha). Environmental conditions were homogeneous for each site.

1.8.7.2 Dispersal Dispersal parameters are species-specific and include maturity age, fecundity (seedling potential per m² crown area), masting (lower seed production in non-mast years), and dispersal distance (probabilistic dispersal kernel). A small fraction of seeds are dispersed over longer distances (long distance dispersal algorithm). Dispersal is calculated at 20 m spatial resolution. Seedling establishment is further modified by other environmental constraints and is stochastically spatially distributed at finer (2 m) resolution. The simulation extent is treated as a spatially explicit landscape, so neighboring cells can serve as seed sources for each other.

We initialized our simulation experiments from bare ground with uniform probability of seed availability set to a low level (0.001) for all species. Once trees reached maturity, they also contributed to seed availability and dispersal.

1.8.7.3 Simulation length Simulations were run for 2010 years, and equilibrium was reached at year 1500. We assessed temporal patterns of species composition and basal area and defined equilibrium as when forest composition and structure stabilized across all sites (“potential natural vegetation”). Some sites had high interannual variation, especially in basal area, after equilibrium; this variation did not represent directional change and appeared instead to be related to sites with extreme growth conditions and variation in randomly selected annual climate.

1.8.7.4 Simulation approaches Only one simulation was run per site.

1.8.8 Output

For each site, we generated decadal outputs for the full simulation extent every 100 years after equilibrium was reached (i.e., 1500-1509, 1600-1609, ..., 2000-2009). This resulted in outputs for a total of 600 ha (100 ha simulation extent x 6 decades). From this list, we randomly sampled 200 ha without replacement.

We derived recruit density and basal area annually by species. Trees > 4 m in height are represented as individuals in iLand, so we were able to identify individual trees that crossed the DBH threshold of either 7 or 10 cm each year. We summed the number of annual recruits and their basal area in the year they crossed the recruitment threshold for each 10-year output period.

We averaged total density and basal area by species for each decade. We further included 0 values (complete cases) for species that were missing from a given sample and site.

1.8.9 References

- Allen, R. G., L. S. Pereira, D. Raes, and M. Smith. 1998. Crop evapotranspiration - Guidelines for computing crop water requirements. FAO - Food and Agriculture Organization of the United Nations, Irrigation and drainage paper 56. Campbell, G. S., and J. M. Norman. 2000. An Introduction to Environmental Biophysics. Springer Science & Business Media. Murray, F. W. 1967. On the computation of saturation vapor pressure. Journal of Applied Meteorology 6:203–204.

1.9 LandClim

1.9.1 Authors

Olalla Díaz-Yáñez, Harald Bugmann

1.9.2 Model

We used LandClim version 2.0, which calculates forest dynamics at a decadal time step. The complete reference resources for the LandClim model and its regeneration module are:

- SCHUMACHER, S., H. BUGMANN, AND D. J. MLADENOFF. (2004). “Improving the formulation of tree growth and succession in a spatially explicit landscape model” Ecological Modelling 180: 175-194.
- SCHUMACHER, S. AND H. BUGMANN. (2006). “The relative importance of climatic effects, wildfires and management for future forest landscape dynamics in the Swiss Alps” Global Change Biology 12: 1435-1450.
- SCHUMACHER, S., B. REINEKING, J. SIBOLD, AND H. BUGMANN. (2006). “Modeling the impact of climate and vegetation on fire regimes in mountain landscapes” Landscape Ecology 21: 539-554.
- Manual available on the model web page: <https://ites-fe.ethz.ch/openaccess/products/landclim>

1.9.3 Climate

Climate inputs for LandClim are at monthly temporal resolution. We used monthly climate from the CHELSA dataset provided in the protocol, using temperature and precipitation variables. During the simulations, we randomly sampled the annual climate data from the CHELSA time series.

1.9.4 Soil

We used the soil quality provided in the protocol to infer the Available Water Holding Capacity (WHC) needed in LandClim. We assigned WHC = 8 cm for soil quality one and WHC = 25 cm for soil quality five, and interpolated the values for the other categories linearly between these two extremes.

1.9.5 Topography

We used the altitude and soil information provided in the protocol to create 200 landscapes, one per site. However, each landscape created per site had no slope, altitude, or aspect variation.

1.9.6 Tree species

We simulated forest dynamics considering only the 11 species proposed in the protocol: *Fagus sylvatica*, *Picea abies*, *Abies alba*, *Carpinus betulus*, *Tilia cordata*, *Acer pseudoplatanus*, *Betula* spp., *Fraxinus excelsior*, *Quercus* spp., *Alnus glutinosa*, and *Pinus sylvestris*.

1.9.7 Simulations

We simulated forest dynamics from bare ground for each site, using patches (pixels) with a size of 20x20 cm². We prepared landscapes of 250x250 m without topographical variation for each site. The simulations we run from bare ground to 2000 years, and the samples were taken starting in the year 2000. In the regeneration module, we used (1) a base seeding probability of 0.001, which defines if seeds are available due to background seed rain; (2) a dispersal mode based on the formulation ‘random asymmetric’; and (3) the same establishment probability of 0.9 for all species included in the simulation of, which sets the same probability for all the species to regenerate successfully, provided that the abiotic and light conditions are suitable for establishment.

1.9.8 Outputs

We estimated the outputs based on a sample of 25 random cells every ten years from the simulation years 2010 to 4000 for one simulation per site, thus providing the required 200 samples per site.

1.10 Landis II

1.10.1 Authors

Josef Brúna, Paola Mairotta, Marco Mina, Giorgio Vacchiano

1.10.2 Model

Landscape scale forest simulation model LANDIS-II (Scheller et al. 2007) with PnET-Succession extension (version 4.1). PnET-Succession is based on the Biomass Succession extension of Scheller

and Mladenoff (2004), embedding elements of the PnET-II ecophysiology model of Aber et al. (1995) to mechanistically simulate growth as a competition for available light and water (De Bruijn et al. 2014). PnET-Succession simulates the competition among cohorts for water and light at a monthly time-step as a function of photosynthetic processes and maintenance respiration that are explicitly linked to environmental drivers such as temperature, precipitation, photosynthetic active radiation (PAR), and CO₂ concentration. Regeneration depends on distance from a seed source, soil water, and subcanopy light, while competition is modeled by partitioning incoming light through multiple canopy layers. Growth of specific cohort biomass components is allocated in the following order: non structural carbon, first, then foliage then root and stem. Biomass growth decreases as cohorts approach their longevity age but mortality can occur at any time when carbon reserves production is insufficient to support growth due to shading, water competition, and/or drought (Gustafson et al. 2015), or by any disturbance or management extension used. For output, we selected Biomass Community Output Extension version 2.0.1, which generates output in the format: species age biomass (g / m²) at specified time steps.

The most complete reference resources for the LANDIS-II model and regeneration module are:

- Scheller R.M. & Mladenoff D.J. (2004) A forest growth and biomass module for a landscape simulation model, LANDIS: design, validation, and application. Ecological Modelling 180: 211–229. <https://doi.org/10.1016/j.ecolmodel.2004.01.022>
- Scheller, R. M., Domingo J. B. , Sturtevant B. R. , Williams J. S. , Rudy A. , Gustafson E. J. & Mladenoff D. J. (2007) Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. Ecological Modelling 201(3-4): 409–419. <https://doi.org/10.1016/j.ecolmodel.2006.10.009>

- de Bruijn A., Gustafson E.J., Sturtevant B.R., Foster J.R., Miranda B.R., Lichti N.I. & Jacobs D.F. (2014) Toward more robust projections of forest landscape dynamics under novel environmental conditions: Embedding PnET within LANDIS-II. Ecological Modelling (287): 44–57 <https://doi.org/10.1016/j.ecolmodel.2014.05.004>
- Petter G., Mairota P., Albrich K., Bebi P., Brůna J., Bugmann H., Haffenden A., Scheller R.M., Schmatz D.R., Seidl R., Speich M., Vacchiano G. & Lischke H. (2020) How robust are future projections of forest landscape dynamics? Insights from a systematic comparison of four forest landscape models. Environmental Modelling & Software 134: 104844. <https://doi.org/10.1016/j.envsoft.2020.104844> (supplementary material)
- Gustafson E.J., De Bruijn A.M.G., Pangle R.E., Limousin J.M., McDowell N.G., Pockman W.T., Sturtevant B.R., Muss J.D. & Kubiske M.E. (2015) Integrating ecophysiology and forest landscape models to improve projections of drought effects under climate change. Global Change Biology 21: 843–856. <https://doi.org/10.1111/gcb.12713>

1.10.3 Climate

Climate input was generated by randomly selecting entire years from the monthly data from the CHELSA dataset and collating them one after the other to create a 2000 years series. The sampling was repeated 10 times with different random order of years to create the 10 climate replicates. Selecting a year at a time eliminates inconsistent seasons (especially summer). We used monthly mean, minimum and maximum temperatures and precipitation as input for PnET succession extension. PnET-Succession also requires mean monthly values of PAR during daylight hours, which

we calculated by dividing the monthly rsds data by the number of daylight seconds in each month for latitude 47 to get W/m². We then multiplied this by 4.6 to get µmol/m²/s and multiplied by 0.5 to get PAR part of the spectrum, based on the approximation by Tsubo et al. (2005). Leap years were not considered. The PAR distribution across all sites and years is in range with values that are typical for temperate regions.

CO₂ concentration with yearly timestep was added from the concentrations table available from EEA (2019). For years 0-1750 CO₂ concentration was maintained fixed to 278 ppm. As for the period 1750-1975 there was only one value each 5 years, we used a constant value for each 5-year period. CO₂ concentration was added to the climate input file after randomization of temperature, precipitation and PAR (see above), so it reflects realistic values for the simulated years. Simulations were run from year 0 to 1990 but sampled outputs did not exceed year 1910. Thus, the drastic increase in CO₂ concentration during the second half of the 1900s did not affect the delivered results.

1.10.4 Soil

Soil type was set to LOAM at all sites. Rooting depth was calculated from soil quality codes (1-5, original values ranged from 1.13 to 4.87) multiplicatively by 205. This produces a soil depth in the range 231 - 998 mm, that are within the typical range for PnET-Succession (maximum 1000 mm). A clear trend emerged - deeper soils resulted in higher biomass, although with some exceptions.

1.10.5 Topography

The ecoregion-specific parameter Precipitation Loss Fraction denotes the water lost to runoff induced by topographic slope. We estimated this parameter for each site from topo_data_200 dataset as slope * 0.01 (e.g., 12% slope = 0.12 PrecLossFrac). We acknowledge that this is a simplification and that better parameterization would be possible if soil types were available for each site. Tests showed that sites with extreme slope tended to have lower biomass, but the signal was not that clear. The correlation between slope and elevation was 0.6 which could obscure the results.

Some sites at very high elevation had, correspondingly, very low temperature. At these sites, the model showed a visible trend in producing lower biomass (down to zero at a few sites). Yet this was probably not apparent in the regeneration results, as extremely low biomass could still support regeneration, even if cohorts do not live long, due to our modification of sexual maturity age.

1.10.6 Tree species

As LANDIS-II cannot start from bare ground, initial communities were set to include all species defined in the protocol at all sites starting with 10-years old cohorts of the following species: fasy (*Fagus sylvatica*), pabi (*Picea abies*), abal (*Abies alba*), cabe (*Carpinus betulus*), tico (*Tilia cordata*), acps (*Acer pseudoplatanus*), betu (*Betula spp.*), frex (*Fraxinus excelsior*), quer (*Quercus spp.*), alg1 (*Alnus glutinosa*), pisy (*Pinus sylvestris*). The species-specific parameter denoting sexual maturity was set to 10 years equally for all species, so that seed source of all species is available from the onset of the simulations at all sites and to maintain seed in the landscape pool. Vegetative regeneration (resprouting) was prevented, since this would have an effect on allometry and regeneration.

Some species may be under or overrepresented because their occurrence in LANDIS-II simulations is typically due to small-scale disturbances or management which were not included in the simulations.

The model requires two sets of parameters representing species-specific life traits. The first set is used by the LANDIS-II core module. The species-specific parameter denoting “longevity” and “sexual maturity” were the only parameters of this set considered in the current exercise. The “sexual maturity” was set to 10 years equally for all species, so that seed source of all species is available from the onset of the simulations at all sites and to maintain seed in the landscape pool. Vegetative regeneration (resprouting) was prevented since this would influence allometry and regeneration.

The second set is used by the PnET-Succession extension and includes several ecophysiological parameters. For this second set we started with the values already applied in Petter et al. (2020) with a previous version of PnET. The values of some of the parameters to which PnET is most sensitive and/or for which the range of variation had been modified for PnET4.1 were then iteratively adjusted (individually and then in combination) to meet the requirements of PnET version 4.1 used in this exercise. The changes were implemented one by one in simulations with the exercise data and those leading to an output deemed realistic and to a relative stabilization of the biomass trend through the simulation were retained for the implementation of the full set of replicates. The retained changes include: 1) decreased percentage of foliar nitrogen for algl (*Alnus glutinosa*); 2) rescaling to the new ranges for all species of the values of the Halfsat and the FracBelowG. The first of these parameters represents shade tolerance and its value affects the probability of establishment of a species (Half saturation light level for photosynthesis.); the latter regulates allocation between

above and below ground pool (Fraction of non-foliar biomass that is belowground (root pool)) ; 3) a slight decrease of the value of the parameter (Fracfol) affecting the yearly allocation to foliage of the active total woody biomass for the three conifers abal (*Abies alba*), piab (*Picea abies*), pisy (*Pinus sylvestris*).

1.10.7 Simulation

Simulation area was defined as a rectangle of 10×20 dimension with sites (i.e., cells) of 1-ha. Each site is an individual ecoregion (i.e., one cell - as in our case - or group of cells sharing similar climate and soil conditions, as a way in LANDIS-II to define differences in biophysical conditions across landscapes) with specific values for soil quality, topography and climate, based on the data provided. The model runs with a yearly timestep using monthly climate. We started each model run with initial communities of all species at age 10 at all sites. Each run was simulated for 1980 years.

1.10.7.1 Dispersal We have used the non-spatial seeding algorithm “universal dispersal” for all the model runs. This way, every species can seed any forest site in the landscape (across all the 200 sites). This mode does not take into account the spatial arrangement and contiguity of cells (i.e., no active seed dispersal), which is proper for the sake of the model comparison exercise. Yet in real landscape seed limitation can also play a role in species composition. The seed number is affected by species-specific age of maturity, which we have modified. This setting ensures an influx of seeds from outside consistent with the original settings.

1.10.7.2 Sampling design We have sampled each site 20 times along the simulation run in two time windows starting 50 years after the first cycle following typical species longevity. After the year 700, biomass of most species was relatively stable and we sampled every 50 years, starting at 750 until the year 1200. The next sampling period started from 1450 and continued until 1900, i.e., after the next longevity peak, again every 50 years. At each selected sampling time, we extracted cohorts of each species according to the age related to the two DBH thresholds (7cm/10cm). Additionally, we extracted cohorts that were up to 9 years older than these thresholds, to get the expected results for a 10-year interval. For example, *Fagus sylvatica* reaches DBH 7cm at 17 years, therefore we have selected all cohorts with age 17-26.

To deliver more than one simulation per site, we created 10 replicates with different climate replicate and random number generator seed number to account for stochastic processes in PnET-Succession. This resulted in the expected 200 samples for each site.

1.10.8 Outputs

Sample number code

Number (1-200) was defined as sample in time + (replicate-1)*20, with sample in time (1-20) according to the sample order. Replicate (1-10) combines random number and climate sequence. For example sample number 25 was taken from the climate replicate 2 and the fifth sample in time (year 950).

r.trees

As LANDIS-II outputs are expressed in terms of cohort age and mean biomass per site, we have calculated the mean DBH and biomass of each species based on age. We estimated the age when

each species reaches 7 cm and 10 cm DBH and the number of individuals using allometric models from the Italian forest inventory collected in R package ForIT (Puletti et al. 2014), and unpublished data from the study by Mina et al. (2021). Please note that the equations used predict DBH >0 already for a tree with age 1, this is because it is based on forestry data and tree cores. The real age when reaching 1.3 m height is species specific and we did not implement this offset. For each cohort, we divided the cohort biomass reported by LANDIS-II, recalculated to kg/ha by the mean biomass of a tree of the same species and age to obtain the number of trees of each species per site. We rounded all results up to the nearest integer to avoid any fraction of trees. The number of individuals was aggregated for all the cohorts of the age of reaching 7 cm or 10 cm DBH and all cohorts of the same species that were up to 9 years older. This approach does not reflect influence of site conditions or stand densities on Age-DBH and Age-Biomass. Relationships between age and DBH were derived by fitting asymptotic regressions using individual tree measurements from an ensemble of forest inventory plots distributed across south-eastern Canada (see Supporting information of Mina et al. 2021). This was the only public dataset for which we could relate tree age, derived from stem coring and tree ring analysis, to measured DBH. Although tree age measurements were obtained with a robust methodology, these functions do not disentangle the influence of site conditions or stand densities. Additionally, they reflect growing conditions of tree genera in Eastern North America.

r.ba basal area of regeneration Calculated as a sum of r.ba of all the cohorts selected in r.trees using the following formula: $DBH^2 * \pi/4 * r.tree$.

ba: total basal area We have not calculated this since our allometric equations do not cover older ages. Therefore, we produced the following two biomass variables as a surrogate. The basal area

was later calculated as a sum of r.ba and estimated ba based on biomass. The estimation includes transformation of kg into tonnes / ha and expansion factor 12.5 to transform ba m²/ha.

((landis\$sum.bio - landis\$r.bio.sum)* 0.001) / 12.5). This includes

sum.bio

Biomass of all cohorts of the species on site including the regeneration and even younger regeneration (units kg/ha).

r.bio.sum Biomass of the regeneration cohorts corresponding to the r. ba and r.trees on site (units kg/ha).

Please notice that we included 0 values for species missing in a given sample and for samples in sites where Landis did not simulate any tree.

1.10.9 References

Aber J. D. & Federer C. A. (1992). A generalized, lumped-parameter model of photosynthesis, evapotranspiration and net primary production in temperate and boreal forest ecosystems. *Oecologia* 92(4): 463-474.

Aber J.D., Ollinger S.V., Federer C.A., Reich P.B., Goulden M.L., Kicklighter D.W., Melillo J.M. & Lathrop R.G. Jr. (1995) Predicting the effects of climate change on water yield and forest production in the northeastern United States. *Climate Research* 5(3): 207–22.

de Bruijn A., Gustafson E.J. , Sturtevant B.R., Foster J.R., Miranda B.R., Lichti N.I. & Jacobs D.F. (2014) Toward more robust projections of forest landscape dynamics under novel environmental conditions: Embedding PnET within LANDIS-II. *Ecological Modelling* (287): 44–57

- EEA (2019) Trends in atmospheric concentrations of CO₂ (ppm), CH₄ (ppb) and N₂O (ppb), between 1800 and 2017 available at https://www.eea.europa.eu/ds_resolveuid/217c026ca03649398aadf39d87623e57
- Gustafson, E. J., A. M. G. De Bruijn, R. E. Pangle, J.-M. Limousin, N. G. McDowell, W. T. Pockman, B. R. Sturtevant, J. D. Muss, and M. E. Kubiske. 2015. Integrating ecophysiology and forest landscape models to improve projections of drought effects under climate change. *Global Change Biology* 21:843-856.
- Mina, M., Messier, C., Duveneck, M., Fortin, M. J., & Aquilué, N. 2021. Network analysis can guide resilience-based management in forest landscapes under global change. *Ecological Applications*, 31(1), 1–18.
- Petter G., Mairotta P., Albrich K., Bebi P., Brúna J., Bugmann H., Haffenden A., Scheller R.M., Schmatz D.R., Seidl R., Speich M., Vacchiano G. & Lischke H. (2020) How robust are future projections of forest landscape dynamics? Insights from a systematic comparison of four forest landscape models. *Environmental Modelling & Software* 134: 104844. <https://doi.org/10.1016/j.envsoft.2020.104844>
- Puletti N., Mura M., Castaldi C., Marchi N., Chiavetta U. and Scotti R. (2014). ForIT: Functions from the 2nd Italian Forest Inventory (INFC). R package version 1.0. <https://CRAN.R-project.org/package=ForIT>
- Scheller, R. M., Domingo J. B. , Sturtevant B. R. , Williams J. S. , Rudy A. , Gustafson E. J. & Mladenoff D. J. (2007) Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecological Modelling* 201(3-4): 409–419. <https://doi.org/10.1016/j.ecolmodel.2006.10.009>

Tsubo M. & Walker S. (2005) Relationships between photosynthetically active radiation and clearness index at Bloemfontein, South Africa. *Theoretical and Applied Climatology* 80: 17–25. <https://doi.org/10.1007/s00704-004-0080-5>

1.11 TreeMig

1.11.1 Authors

Heike Lischke

1.11.2 Model

The model TreeMig (see treemig.wsl.ch) was originally based on the ForClim model, then aggregated by a distribution based approach (Lischke et al. 1998), and extended to a forest landscape model by including seed production, seed dispersal, seed bank dynamics and seedling establishment (Lischke and Loffler 2006, Lischke et al. 2006).

The most complete reference resource for the TreeMig model and regeneration module are:

- Lischke, H., T. J. Loffler, and A. Fischlin. 1998. Aggregation of individual trees and patches in forest succession models: Capturing variability with height structured, random, spatial distributions. *Theor Popul Biol* 54:213-226.
- Lischke, H., N. E. Zimmermann, J. Bolliger, S. Rickebusch, and T. J. Loffler. 2006. TreeMig: A forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. *Ecological Modelling* 199:409-420.

- Lischke, H., and T. J. Loffler. 2006. Intra-specific density dependence is required to maintain species diversity in spatio-temporal forest simulations with reproduction. *Ecological Modelling* 198:341-361.
- The online resources available at treemig.wsl.ch

1.11.3 Climate

As climatic input, the monthly time series of mean temperature and precipitation sum were used. With these climatic variables and additional data about slope, aspect and soil field capacity, yearly time series of the following bioclimatic variables were calculated: (1) DDSum, the yearly day degree sum above 5.5°C , (2) MinWiT, the lowest mean monthly temperature, (3) DrStr, drought stress according to a bucket model that estimates a monthly PET following (Thornthwaite and Mather 1957), using temperature, latitude, slope and aspect as input, and calculates the monthly soil water content based on soil water content in the previous month, precipitation, interception, and field capacity (“bucketsize”) (Bugmann 1994, Fischlin et al. 1995). Drought stress is then given by 1- the ratio between demand (PET-interception) and the supply, a function of the soil water content. To get climate data also for the spin-up, bioclimate was calculated from the monthly data as given, and then extended by sampling from the given data from the end of the climate data to get 1000 years. The same sequence of sampling years was used in all simulations and sites.

1.11.4 Soil

The soil bucketsize (available water capacity of the upmost 1m soil layer) required for the drought stress calculation, was derived from the soil quality data set, by $\text{Bucketsize} = 10 + (\text{soil_quality}-1)$

/4 * 25 , to covert the range from 10 cm to 35cm

1.11.5 Topography

Slope and aspect were taken from the topographic data set and converted to a common variable entering the PET calculation, by $\text{slasp} = 2\cos(\text{aspect} \pi/180) * \min(1, \text{slope}/60)$

1.11.6 Tree species

Thirty Central European tree species were simulated, including the standard species of the project. Only *Betula pubescens* was not simulated, because we did not have a parametrization for this species at hand. Species: *Abies alba*, *Larix decidua*, *Picea abies*, *Pinus cembra*, *Pinus montana*, *Pinus sylvestris*, *Taxus baccata*, *Acer campestre*, *Acer platanoides*, *Acer pseudoplatanus*, *Alnus glutinosa*, *Alnus incana*, *Alnus viridis*, *Betula pendula*, *Carpinus betulus*, *Castanea sativa*, *Corylus avellana*, *Fagus sylvatica*, *Fraxinus excelsior*, *Populus nigra*, *Populus tremula*, *Quercus petraea*, *Quercus pubescens*, *Quercus robur*, *Salix alba*, *Sorbus aria*, *Sorbus aucuparia*, *Tilia cordata*, *Tilia platyphyllos*, *Ulmus glabra*

1.11.7 Simulation

The simulations were carried out on a grid of 15 x 15 cells, each with 200 m side length.

1.11.7.1 Dispersal The seeds produced by the mother trees are transported from the source cell according to a double negative exponential, species specific that determines the share of the seeds in the source cell landing in a sink cell in a given distance to the source cell. In the sink cells,

the seeds enter the seed bank . At the start of the simulation for 10 years seeds of all species are available in all grid cells, from then on, the normal seed dispersal starts.

1.11.7.2 Regeneration processes TreeMig simulates the full feedback from seed production to ingrowth. Some of the involved processes depend on environment. (1) Overall seed production in TreeMig is described by multiplying a species specific seed production per tree that depends on tree size and thus indirectly on the growth, which depends on DDsum, drought stress, nutrients and light and the adult tree number that depends on survival that depends on DDsum, drought stress, nutrients and light. (2) Seed dispersal is independent of environment. (3) Seed bank dynamics – also environment independent - is given by seed input, seed mortality, and germination. Additionally, the number of seeds S of a species in the seed bank of this species (SB) is limited to a carrying capacity C of 1000, by

$$SB = C + (SB - C) * \text{Exp}(-S/C).$$

(4) Germination and seedling survival depend on browsing, drought, winter temperature, DDSum and light. (5) The saplings up to 1.37 die and grow like adult trees, depending on DDsum, drought stress and light. Only the parameter value of the light dependence differs from that of the adult trees.

The regeneration processes were simulated annually in the order 1. germination, 2. germinated seeds leave the seed bank, 3. mortality of seeds in the seed bank, 4. seed entry by seed dispersal. The ingrowth numbers were very sensitive to this order. Simulations with another sequence, in which the seeds in the seed bank died first and then the remaining ones germinated, resulted in about half the ingrowth.

1.11.7.3 Simulation length The simulations were started on bare ground and run for 1000 years, when at all sites the equilibrium was reached.

1.11.7.4 More simulations per site TreeMig simulates directly the mean dynamics of entire stands, but based on the variability of light conditions within the stand (distribution based approach, cf. Lischke et al. 1998). To come up with the required 200 random patches of 1 ha, temporal and spatial variability intrinsic to TreeMig were combined. Temporal variability was given by sampling from each decade of the last 100 years of the simulation, i.e. years 900-1000. For each of these 10 decades, 20 grid cells were sampled from the simulated grid, reflecting spatial variability, results from demographic stochasticity and short-range spatial interactions

1.11.8 Outputs

In contrast to gap models, TreeMig is based on the concept of frequency distributions of tree densities classes in different tree heights on ca. 1/12 ha (833m^2) patches, and the resulting frequency distributions in discrete light. During the simulation, in each height class the frequencies of the different light classes were recorded, as well as the ingrowth of each species in the height-light class. By sampling 12 times from the height-light class ingrowths according to the height-light class frequency distribution and summing up, the stochastic ingrowths into all TreeMig height classes for one ha were calculated for each grid cell. Then the height class boundaries were translated into DBH and the ingrowths into the height classes linearly interpolated to 7cm and 10cm. The ingrowth basal areas into the height classes were obtained by multiplying the interpolated ingrowth numbers with the basal areas at 7cm and 10cm, i.e. $r.\text{trees} * \text{Pi} * (0.07/2)^2$ and $r.\text{trees} * \text{Pi} * (0.1/2)^2$. The species specific basal area (at 1.37m) sums are a standard output of TreeMig, given by multi-

plying the state variables (number of trees per grid cell, height class, and species) with the species specific basal area of each height class. To account for the ingrown trees dying within a decade, as an approximation the survival in each year to the power of 5 was used. Thereby, it was assumed that the trees have to survive in average 5 years in each decade, and that the current mortality is valid for these 5 years.

1.11.9 References

- Bugmann, H. 1994. On the ecology of mountainous forests in a changing climate: A simulation study. Dissertation. Swiss Federal Institute of Technology, Zurich.
- Fischlin, A., H. Bugmann, and D. Gyalistras. 1995. Sensitivity of a forest ecosystem model to climate parametrization schemes. Env. Poll. 87:267-282.
- Lischke, H., and T. J. Loffler. 2006. Intra-specific density dependence is required to maintain species diversity in spatio-temporal forest simulations with reproduction. Ecological Modelling 198:341-361.
- Lischke, H., T. J. Loffler, and A. Fischlin. 1998. Aggregation of individual trees and patches in forest succession models: Capturing variability with height structured, random, spatial distributions. Theor Popul Biol 54:213-226.
- Lischke, H., N. E. Zimmermann, J. Bolliger, S. Rickebusch, and T. J. Loffler. 2006. TreeMig: A forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. Ecological Modelling 199:409-420.
- Thornthwaite, C. W., and J. R. Mather. 1957. The water balance. Publications in Climatology

VIII:1-69.

1.12 LPJ-GUESS

1.12.1 Authors

Tim Anders, Jessica Hetzer, Thomas Hickler

1.12.2 Model

The model was originally developed by Ben Smith of Lund University in a collaboration also involving the Potsdam Institute for Climate Impact Research and the Max-Planck Institute for Biogeochemistry. Over the years, many people from institutes around the world have contributed to the refinement and further development of the model.

We used the process-based mechanistic model LPJ-GUESS to simulate vegetation dynamics considering tree species, age cohorts, gap dynamics and biogeochemical cycles. Parameters of the model reflect an updated version of the most common European tree species as well as typical shrub PFTs as described by Hickler et al. (2012).

The most complete reference resource for the LPJ-GUESS model and regeneration module is:

- Smith, B. (2001). LPJ-GUESS-an ecosystem modelling framework. Department of Physical Geography and Ecosystems Analysis, INES, Sölvegatan, 12, 22362.
- Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production

in an individual-based dynamic vegetation model. Biogeosciences, 11(7), 2027-2054.

- Source code is available on demand (see <https://web.nateko.lu.se/lpj-guess/>).

1.12.3 Climate

We used the provided daily time series of climatic variables from 1981 to 2018. The climatic variables base on the CHELSA data set with a spatial resolution of $0.008^\circ \times 0.008^\circ$ (~ 1 km) (Karger et al., 2021). Daily average temperature (tas) [$^\circ\text{C}$], daily average maximum temperature (tasmax) [$^\circ\text{C}$], daily average minimum temperature (tasmin) [$^\circ\text{C}$], precipitation sum (pr) [mm] and surface solar (shortwave) radiation (rsds) were taken to set up the LPJ-GUESS simulation. Used climatic variables were converted to netCDF files. As LPJ-GUESS requires the climatic variables in specific units, temperature variables (tas, tasmax, tasmin) [$^\circ\text{C}$] were converted to Kelvin and surface solar radiation (rsds) [Jm^{-2}] was converted to W/m^2 by dividing rsds by 86400 s (3600 s/h * 24 h). Nitrogen deposition was considered to be constant at a level of 10 kgN/ha/year . Atmospheric carbon dioxide concentration data from 1900 to 2018 was taken from the global carbon project (Quéré et al. 2018).

1.12.4 Soil

All sites were simulated with default values for a medium textured soil type (sand: 0.35, clay: 0.15, 823 silt: 0.5) in LPJ-GUESS.

1.12.5 Topography

Topography was not considered.

1.12.6 Tree species

LPJ-GUESS can be run in different modes, with different levels of abstraction of the population and community processes. For this study, the more detailed ‘cohort’ mode was used, in which individuals, patches and vertical canopy structure are represented explicitly, but living individuals within a cohort (age class) of a given tree species in a given patch are assumed to be identical (in terms of all state variables, such as height and stem diameter). Simulations consider the main tree species as well as other PFTs yielding an appropriate representation of European vegetation, including only tree and shrub species that are widely distributed across Europe and can become dominant in some areas (Hickler et al., 2012). Mediterranean rain-green small shrubs with shallow roots (e.g. *Lavendula* spp., *Rosmarinus* spp.) and alpine/arctic shrubs (e.g. *Vaccinium* spp.) were represented as PFTs in the model. As in earlier applications of LPJ-GUESS, herbaceous vegetation was represented by two ‘generic herb’ PFTs, with C3 and C4 photosynthesis, respectively. The final set included 16 tree species, one Mediterranean shrub PFT, one boreal/alpine shrub PFT, and the two herbaceous PFTs described earlier (Table 2). To increase the chances of shade-intolerant species becoming established, patch-destroying disturbances were allowed as described in Hickler et al. 2012. *Acer pseudoplatanus* and *Alnus glutinosa* are not parameterized in LPJ-GUESS and therefore not considered.

1.12.7 Simulation

In this study, vegetation is simulated in so called patches that represent forest areas of fixed size. Patches can be regarded as sample stands in a grid cell, each of which represents an 0.1 ha (comparable to figure 1 of the tree regeneration workshop protocol). In this study, vegetation dynamics

were simulated by 2000 replicate patches each 0.1 ha in size. To match the required sample size, we processed the data so that ten 0.1 ha patches were merged to one 1 ha sample.

1.12.7.1 Dispersal Model formulations of establishment are based on those employed within the ‘forest gap’ model FORSKA (Leemans & Prentice, 1989; Prentice et al., 1993). The number of new saplings of each woody PFT/species and in each patch each year is drawn at random from the Poisson distribution, with an expectation influenced by a PFT/species-specific maximum establishment rate and by the ‘propagule pool’, i.e. the amount of carbon allocated to reproduction by all individuals of the PFT/species at all patches in the previous year. No saplings are established in a given patch if the minimum PAR level at the forest floor is below a PFT/species-specific threshold, which is higher for more light-demanding PFTs/species.

1.12.7.2 Simulation length What is the simulation length, and how have you determined the simulation time until the equilibrium?

The simulations were initialized from bare ground (no biomass) and the model was spun up for 500 years until the modelled vegetation was in approximate equilibrium with the climate and CO₂ around the year 1900 (Smith et al., 2001; Hickler et al., 2012).

If you have run more than one simulation per site, please describe the differences of the approaches used. We performed one simulation per site.

1.12.8 Output

The generated output file is called Output_SGN_LPJGUESS.txt. Each row of the output dataset contains variables regarding the decade 2008-2018. The first two columns describe the study site

(column “site”, values reflect site ID) and the 1 ha samples of that specific site (column “sample”, values range from 1 to 200). Output variables are species-specific. Although we simulated a total of 20 European tree species and PFTs, our output file considers only the species required by the protocol (column “species”, values are the abbreviations for the species name (“abal” for *Abies alba*, “cabe” for *Carpinus betulus*, “fasy” for *Fagus sylvatica*, “frex” for *Fraxinus excelsior*, “pabi” for *Picea abies*, “pisy” for *Pinus sylvestris*, “tico” for *Tilia cordata*, “betu_pen” for *Betula pendula*, “betu_pub” for *Betula pubescens*, “quer_il” for *Quercus ilex*, “quer_pub” for *Quercus pubescens*, and “quer_rob” for *Quercus robur*). Since *Acer pseudoplatanus* (“acps”) and *Alnus glutinosa* (“algl”) are not parameterized for LPJ-GUESS so far, we have not included these species in the output. Further we excluded shrubs (alpine/arctic shrubs and *Quercus coccifera*) and herbaceous vegetation from the final output as this study focusses on tree regeneration only. The first variable is the number of 7 cm recruits “r_7cm.trees” calculated as the sum of all trees of the same species that cross the threshold of 7 cm DBH between 2008 and 2018 in that specific sample of that specific site. The variable “r_10cm.trees” refers to 10 cm threshold respectively. The 7 cm recruits’ basal area „r_7cm.ba” is calculated as a weighted sum of r_7cm.trees multiplied by recruits basal area at that time (with $ba=(dbh/2)\cdot\pi$). Similarly, r_10cm.ba was calculated for 10 cm recruits. The last two columns describes the mean annual basal area and the mean number of trees per site, sample, and species overall years between 2008 and 2018 (column “ba”).

Table 1: Species/PFT characteristics and parameter of Hickler et al. 2012. r base respiration rate, kallm1 constant in allometry equations, Tc,min minimum winter temperature for establishment, Tc,max maximum coldest-month temperature for establishment, GDD5 minimum degree-day sum above 5 °C for establishment, fAWC minimum growing-season fraction of available soil

water holding capacity in the first soil layer, rfire fraction of individuals surviving fire, k_{leaf}:sa leaf longevity, CA_{max} maximum woody crown area. Parameter that are associated with establishment processes are marked in bold. Species/PFT

Species/PFT	Short name	Geographic range	r^* (gCgN 'day')	Growth form	k_{allow}^* (°C)	$T_{c,\min}^*$ (°C)	$T_{c,\max}^*$ (°C)	GDD_5^* (°Cd)	fAWC	Chilling requirement*	Shade tolerance class	r_{fire}	$k_{\text{leaf}}:sa^*$	Z_t^{**} (yr)	a_{leaf}^* (yr)	a_{end}^* (yr)	CA_{\max}^b (m ²)
<i>Abies alba</i>	abal	temperate	0,055	tree	150	-4,5	-2	1450	0,35	-	tolerant	0,1	4000	0,8	4	350	40
<i>Betula pendula</i>	betu_pen	temperate	0,055	tree	250	-30	-	700	0,42	intermediate	intolerant	0,1	5000	0,8	0,5	200	40
<i>Betula pubescens</i>	betu_pub	boreal	0,11	tree	250	-	-	350	0,5	intermediate	intolerant	0,1	5000	0,8	0,5	200	40
<i>Carpinus betulus</i>	cabe	temperate	0,055	tree	250	-8	-	1200	0,33	high ^{ab}	intermediate	0,1	5000	0,7	0,5	350	40
<i>Corylus avellana</i>	cor_ave	temperate	0,055	tree	250	-8	-	800	0,3	intermediate	intolerant	0,1	4000	0,7	0,5	300	15
<i>Fagus sylvatica</i>	fasy	temperate	0,055	tree	250	-3,5	-	1500	0,3	high ^{ab}	tolerant	0,1	5000	0,8	0,5	500	40
<i>Fraxinus excelsior</i>	frex	temperate	0,055	tree	250	-16	-	1100	0,4	low	intermediate	0,1	5000	0,8	0,5	350	40
<i>Juniperus oxycedrus</i>	jun_oxy	Mediterranean	0,055	shrub	150	0	-	2200	0,01	-	intolerant	0,4	1500	0,5	1,5	200	10
<i>Picea abies</i>	pabi	boreal	0,11	tree	150	-30	-1,5	600	0,43	-	tolerant	0,1	4000	0,8	4	500	40
<i>Pinus halepensis</i>	pin_hal	Mediterranean	0,055	tree	150	3	-	3000	0,05	-	intolerant	0,4	2000	0,6	2	350	40
<i>Pinus sylvestris</i>	pisy	boreal	0,11	tree	150	-30	-1	500	0,25	-	intermediate	0,4	2000	0,6	2	500	40
<i>Quercus coccifera</i>	quer_coc	Mediterranean	0,055	shrub	250	0	-	2200	0,1	-	intermediate	0,3	2500	0,5	1,5	350	10
<i>Quercus ilex</i>	quer_il	Mediterranean	0,055	tree	250	-2	-	1800	0,1	-	intermediate	0,3	3000	0,5	2	350	40
<i>Quercus pubescens</i>	quer_pub	supra-Mediterranean	0,055	tree	250	-5	-	1900	0,2	low	intermediate	0,2	4000	0,6	0,5	500	40
<i>Quercus robur</i>	quer_rob	temperate	0,055	tree	250	-16	-	1100	0,25	low	intermediate	0,2	4000	0,6	0,5	500	40
<i>Tilia cordata</i>	tico	temperate	0,055	tree	250	-18	-	1000	0,33	high ^{ab}	intermediate	0,1	5000	0,8	0,5	350	40
Boreal evergreen shrub ^e	BES	boreal	0,11	shrub	250	-	-1	300	0,25	-	intolerant	0,1	500	0,8	2	50	3
Mediterranean raingreen shrub ^e	MRS	Mediterranean	0,055	shrub	250	0	-	2200	0,01	low	intolerant	0,3	1500	0,9	1	100	10
C ₃ herbaceous		temperate-boreal	0,055	herbaceous	-	-	-	-	0,01	-	-	0,5	-	0,9	1	-	-
C ₄ herbaceous		tropical	0,011	herbaceous	-	15,5	-	-	0,01	-	-	0,5	-	0,9	1	-	-

1.12.9 References

- Hickler, T., Vohland, K., Feehan, J., Miller, P. A., Smith, B., Costa, L., ... & Sykes, M. T. (2012). Projecting the future distribution of European potential natural vegetation zones with a generalized, tree species-based dynamic vegetation model. *Global Ecology and Biogeography*, 21(1), 50-63.
- Karger, D. N., Lange, S., Hari, C., Reyer, C. P., & Zimmermann, N. E. (2021). CHELSA-W5E5 v1. 1: W5E5 v1. 0 downscaled with CHELSA v2. 0.
- Leemans, R., & Prentice, I. C. (1989). FORSKA-a general forest succession model. *Meddelanden*

Från Växtbiologiska Institutionen. Prentice et al., 1993

Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Hauck, J., Pongratz, J., ... & Zheng, B. (2018). Global carbon budget 2018. *Earth System Science Data*.

Smith, B. (2001). LPJ-GUESS-an ecosystem modelling framework. *Department of Physical Geography and Ecosystems Analysis, INES, Sövegatan, 12, 22362.*

Smith, B., Wårlind, D., Arneth, A., Hickler, T., Leadley, P., Siltberg, J., & Zaehle, S. (2014). Implications of incorporating N cycling and N limitations on primary production in an individual-based dynamic vegetation model. *Biogeosciences*, 11(7), 2027-2054.

1.13 aDGVM2

1.13.1 Authors

Simon Scheiter Senckenberg Biodiversity and Climate Research Centre, Senckenberganlage 25, 60325 Frankfurt am Main, Germany

1.13.2 Model

We used aDGVM2, an individual-based dynamic vegetation model that is based on concepts from community assembly theory and uses a functional trait approach (Langan et al., 2017; Scheiter et al., 2013). The model simulates growth, reproduction, and mortality of individual plants while keeping track of state variables, such as biomass, height, and leaf area. Each plant in aDGVM2 is characterized by a plant-specific set of trait values. Traits describe growth form, leaf characteristics, hydraulic characteristics, resource allocation, architecture, reproduction, mortality, and

response to disturbance. Most plant traits are linked by trade-offs to constrain possible trait combinations. Selection and trait inheritance assemble plant communities that are adapted to biotic and abiotic conditions. Plants with trait combinations that allow sufficient growth and reproduction rates, and that allow plants to cope with competition and disturbances can contribute their trait values to the community trait pool. Trait mutation and recombination may alter trait values in the community trait pool. Randomly drawn seeds from the trait pool are added to the plant population and new plants start growing from seeds. Plants that are not adapted to the prevailing disturbance regimes, biotic and abiotic conditions, or that do not allocate enough carbon to reproduction disappear from the population. Therefore, successful ecological strategies emerge dynamically from these community assembly and reproduction process. Due to the trait-based approach, species or plant functional types (PFTs) are not pre-defined and hard-coded in aDGVM2. Rather, different PFTs or ecological strategies that are adjusted to the prevailing biotic and abiotic conditions emerge from community assembly processes (Scheiter et al., 2013), and plants can be classified into PFTs or ecological strategies based on their trait values in a post-processing step. The aDGVM2 has primarily been developed for tropical and sub-tropical ecosystems and it has not been tested and benchmarked for Europe. We applied the model in the original version. Data-model comparisons with model results were not done and the model was not changed to improve agreement with any data for Europe.

The most complete reference resource for the aDGVM2 model and regeneration module is:

- Scheiter S, Langan L, Higgins SI (2013) Next generation dynamic global vegetation models: learning from community ecology. *NEW PHYTOLOGIST*, 198, 957-969.

1.13.3 Climate

The aDGVM2 simulates vegetation at daily temporal resolution. We therefore used daily climate input data for the period 1981-2018 according to the modeling protocol.

1.13.4 Soil

Soil quality data was converted to maximum rooting depth. Specifically, the soil quality value multiplied by 2 and the result was used as maximum rooting depth. We used the multiplication, as we typically use maximum soil depth and maximum rooting depth of 10m. Previous model simulations showed that rooting depth and soil dept have strong impacts on vegetation dynamics by influencing water availability to plants (Langan et al., 2017). We therefore considered it as adequate representation of soil quality. Other soil characteristics including volumetric water-holding capacity, soil hydraulic conductivity, soil bulk density, soil depth, soil texture, soil carbon content, soil wilting point and field capacity were taken from the FAO (<http://www.fao.org>, Nachtergael et al., 2009).

1.13.5 Topography

We used elevation according to the modeling protocol. In aDGVM2, elevation is used to calculate atmospheric pressure which in turn influences ecophysiological processes related to water and carbon fluxes at leaf and canopy level. Information on slope and aspect was not used, as related processes (e.g., impacts on runoff or radiation balance) are not implemented in aDGVM2.

1.13.6 Tree species

The aDGVM2 does not simulate pre-defined species or PFTs. The 11 tree species included in the modeling protocol cannot be implemented and parametrized with reasonable effort. We therefore classified simulated trees into different ecological strategies as a post-processing step. We used all combinations of evergreen/deciduous, light/water triggered phenology, low/high SLA. Evergreen/deciduous phenology and light/water triggered phenology are traits implemented in aDGVM2. The threshold to distinguish high and low SLA was defined as 20 m²/kg. We used these traits as previous simulation experiments showed that patterns of these traits emerge along environmental gradients while patterns of other traits implemented in aDGVM2 are often less clear. Ecological strategies considered are: low SLA, deciduous, rain-triggered; low SLA, deciduous, light-triggered; low SLA, evergreen, rain-triggered; low SLA, evergreen, light-triggered; high SLA, deciduous, rain-triggered; high SLA, deciduous, light-triggered; high SLA, evergreen, rain-triggered; high SLA, evergreen, light-triggered.

1.13.7 Simulations

Simulations followed the modeling protocol. Simulations were conducted for all sites using the provided daily climate and soil data. By default, aDGVM2 simulates 1ha stands. To implement the modeling protocol, adjustments to scale model results to 1 ha level were therefore not required. The model was initialized by default initialization routines, i.e., traits of individual plants were drawn randomly from a pre-defined range of values for each trait. Ranges of trait values are provided in Langan et al. (2017). The model was initialized with bare ground and 2880 tree seedlings with 100g each. The aDGVM2 simulates both trees and grasses. As the simulation experiment aims at

tree/forest regeneration, simulations were conducted only with trees while grasses were removed.

We conducted 5 replicates for each of the 200 sites to account for stochastic processes in aDGVM2.

Stochastic processes include random initialization of trait values of initial plant populations, demographic processes such as plant mortality or selection of seeds from seed bank, or the community assembly processes including mutation and cross-over. Climate time series used for different replicates were identical such that variation between replicates can be attributed by stochastic processes in aDGVM2.

Although aDGVM2 includes routines to simulate fire, simulations were conducted in the absence of fire. Management was not simulated.

1.13.8 Dispersal

The aDGVM2 does not simulate seed dispersal.

1.13.9 Simulation length

We first conducted a 600-year spin-up with a randomized order of years 1981-2018 provided in the CHELSA daily climate data. Spin-up is required to ensure that both state variables such as biomass and tree cover and community trait composition have stabilized and are in a dynamic equilibrium with climate. Previous model simulations showed that a 600-year spin-up is sufficient to reach such an equilibrium state (Langan et al., 2017). Following the spin-up, we simulated transient vegetation dynamics with daily CHELSA climate data for the period 1981-2018.

1.13.10 Outputs

To create output variables, we randomly selected 10-year periods within the transient phase (i.e., period 1981-2018 after spin-up phase) and calculated averages of all required output variables following the modeling protocol. The selected 10-year periods differed between sites and replicates.

The aDGVM2 does not simulate pre-defined species or PFTs and trees were classified into different ecological strategies (see section Tree species). Codes for the ecological strategies in the species column of the output files are: 1 - low SLA, deciduous, rain-triggered 2 - low SLA, deciduous, light-triggered 3 - low SLA, evergreen, rain-triggered 4 - low SLA, evergreen, light-triggered 5 - high SLA, deciduous, rain-triggered 6 - high SLA, deciduous, light-triggered 7 - high SLA, evergreen, rain-triggered 8 - high SLA, evergreen, light-triggered

Simulations were conducted for all 200 sites. We conducted 5 replicates per site to account for stochastic processes in aDGVM. In the output files, replicates are given as follows: Replicate 1 – sample 1:200 Replicate 2 – sample 201:400 Replicate 3 – sample 401:600 Replicate 4 – sample 601:800 Replicate 5 – sample 801:1000

1.13.11 References

Langan L; Higgins SI; Scheiter S (2017) Climate-biomes, pedo-biomes or pyro-biomes: which world view explains the tropical forest - savanna boundary in South America? JOURNAL OF BIOGEOGRAPHY, 44, 2319-2330.

Nachtergael, F., van Velthuizen, H., Verelst, L., Batjes, N., Dijk- shoorn, K., Van Engelen, V., Fischer, G., Jones, A., Montanarella, L., and Petri, M.: Harmonized world soil database (version

1.1), FAO, Rome, Italy, IIASA, Laxenburg, Austria, 2009.

Scheiter S, Langan L, Higgins SI (2013) Next generation dynamic global vegetation models: learning from community ecology. *NEW PHYTOLOGIST*, 198, 957-969.

For Review Only

2 Supplementary figures and tables

2.1 Ingrowth levels, tree diversity and mortality in tree establishment

Table S1: Number of total samples, samples with zero recruitment and the proportion of zero recruitment per model and observed data.

Model	Total (7 cm)	Total (10cm)	No recruitment (7 cm)	No recruitment (10cm)	Percentage of no recruitment (7 cm)	Percentage of no recruitment (10 cm)
Observed	5676	6564	203	238	3.63	3.58
4C	133000	165000	23898	30177	18.29	17.97
ForCEEPS	363000	440000	1	137	0.03	0.00
ForCEEPS(f)	363000	440000	42	202	0.05	0.01
FORMIND	363000	440000	8663	10082	2.29	2.39
ForClim 1	396000	480000	7	28	0.01	0.00
ForClim 11	396000	480000	77	192	0.04	0.02
SIBYLA	363000	440000	1565	5539	1.26	0.43
xComp	297000	360000	0	0	0.00	0.00
PICUS	363000	440000	400	400	0.09	0.11
iLand	363000	440000	26	143	0.03	0.01
LandClim	363000	440000	378	1347	0.31	0.10
Landis II	363000	440000	28621	34844	7.92	7.88
TreeMig	396000	480000	0	0	0.00	0.00
LPJ-GUESS	328000	398000	6742	7140	1.79	2.06
aDGVM2	6600	8000	603	824	10.30	9.14

Table S2: Test of the significance of the differences of Shannon index in each model for 7 and 10 cm threshold.

model	y	Group 1	Group 2	p	p.signif
Observed	ShannonIndexRecruit	7	10	0.5752300	ns
4C	ShannonIndexRecruit	7	10	0.0112400	*
ForCEEPS	ShannonIndexRecruit	7	10	0.0022400	**
ForCEEPS(f)	ShannonIndexRecruit	7	10	0.0001500	***
FORMIND	ShannonIndexRecruit	7	10	0.0018200	**
ForClim 1	ShannonIndexRecruit	7	10	0.0000000	****
ForClim 11	ShannonIndexRecruit	7	10	0.0000078	****
SIBYLA	ShannonIndexRecruit	7	10	0.0017100	**
xComp	ShannonIndexRecruit	7	10	0.0089900	**
PICUS	ShannonIndexRecruit	7	10	0.0343300	*
iLand	ShannonIndexRecruit	7	10	0.0539100	ns
LandClim	ShannonIndexRecruit	7	10	0.0000000	****
Landis II	ShannonIndexRecruit	7	10	0.3487000	ns
TreeMig	ShannonIndexRecruit	7	10	0.4182800	ns
LPJ-GUESS	ShannonIndexRecruit	7	10	0.8213200	ns

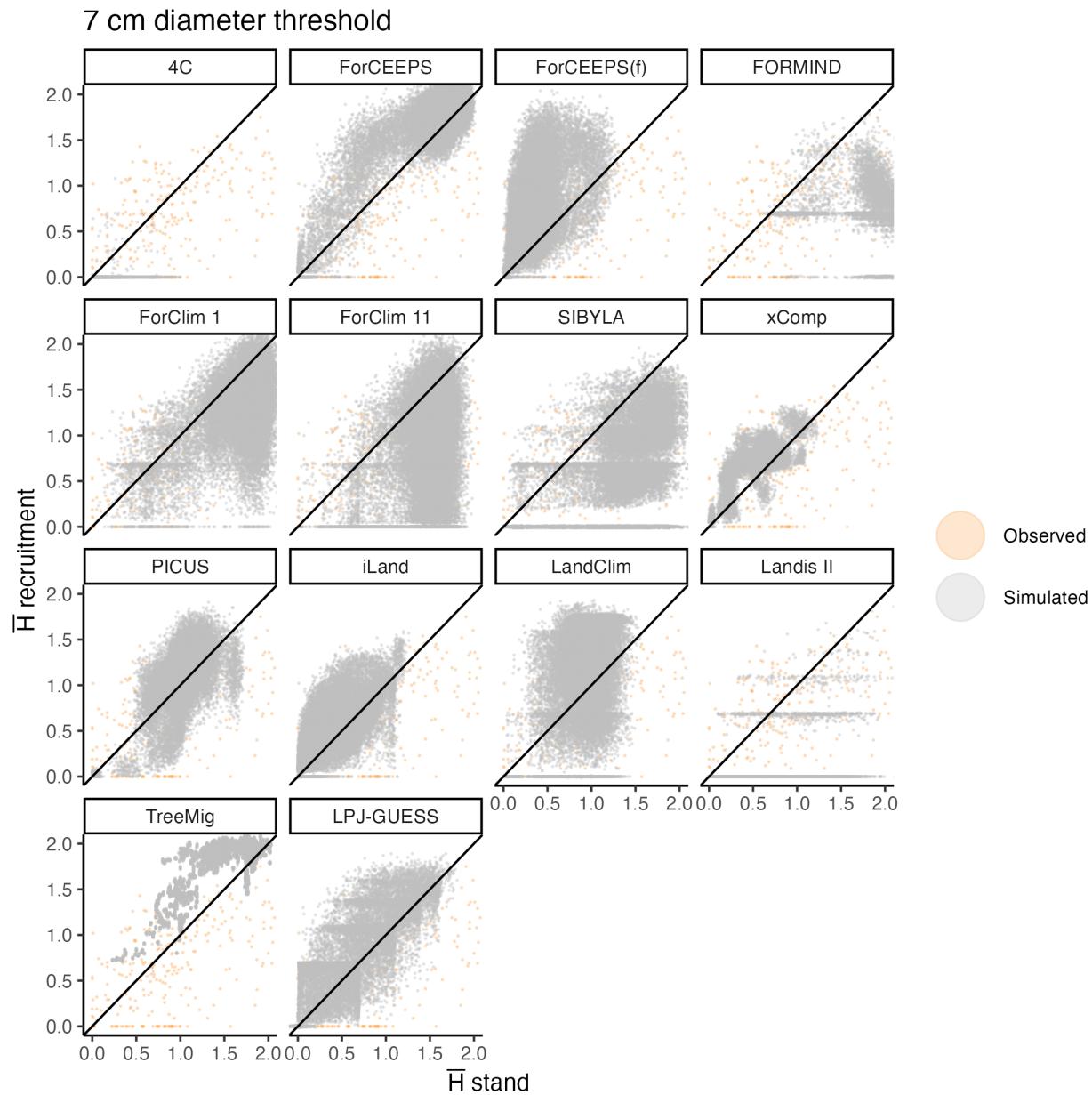


Figure S1: Mean Shannon index per site at recruitment and at the stand level both for observed and simulated values for the diameter threshold of 7cm. The black line marks where the mean Shannon index per site at recruitment and at the stand level would be equal.

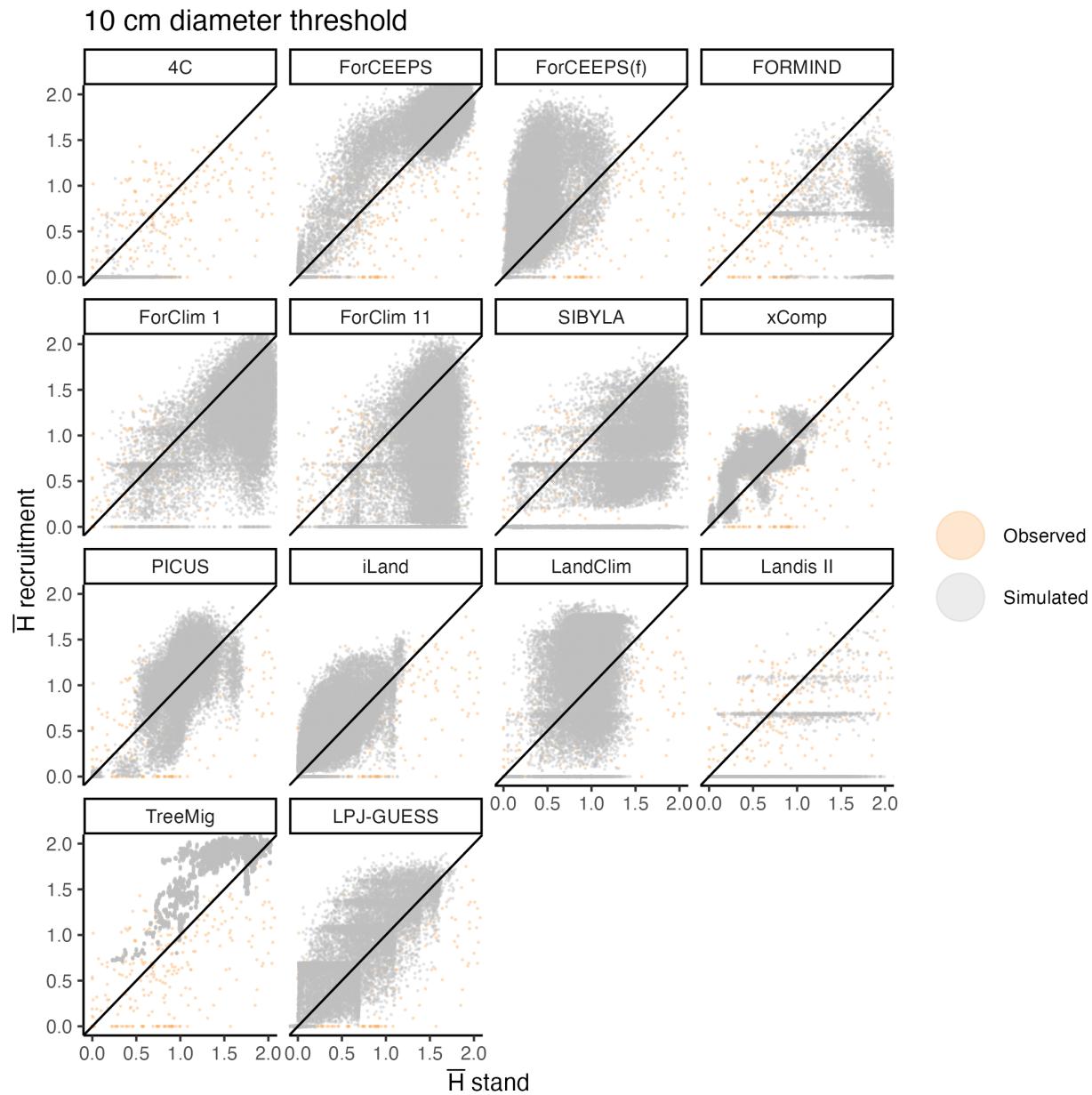


Figure S2: Mean Shannon index per site at recruitment and at the stand level both for observed and simulated values for the diameter threshold of 10cm.

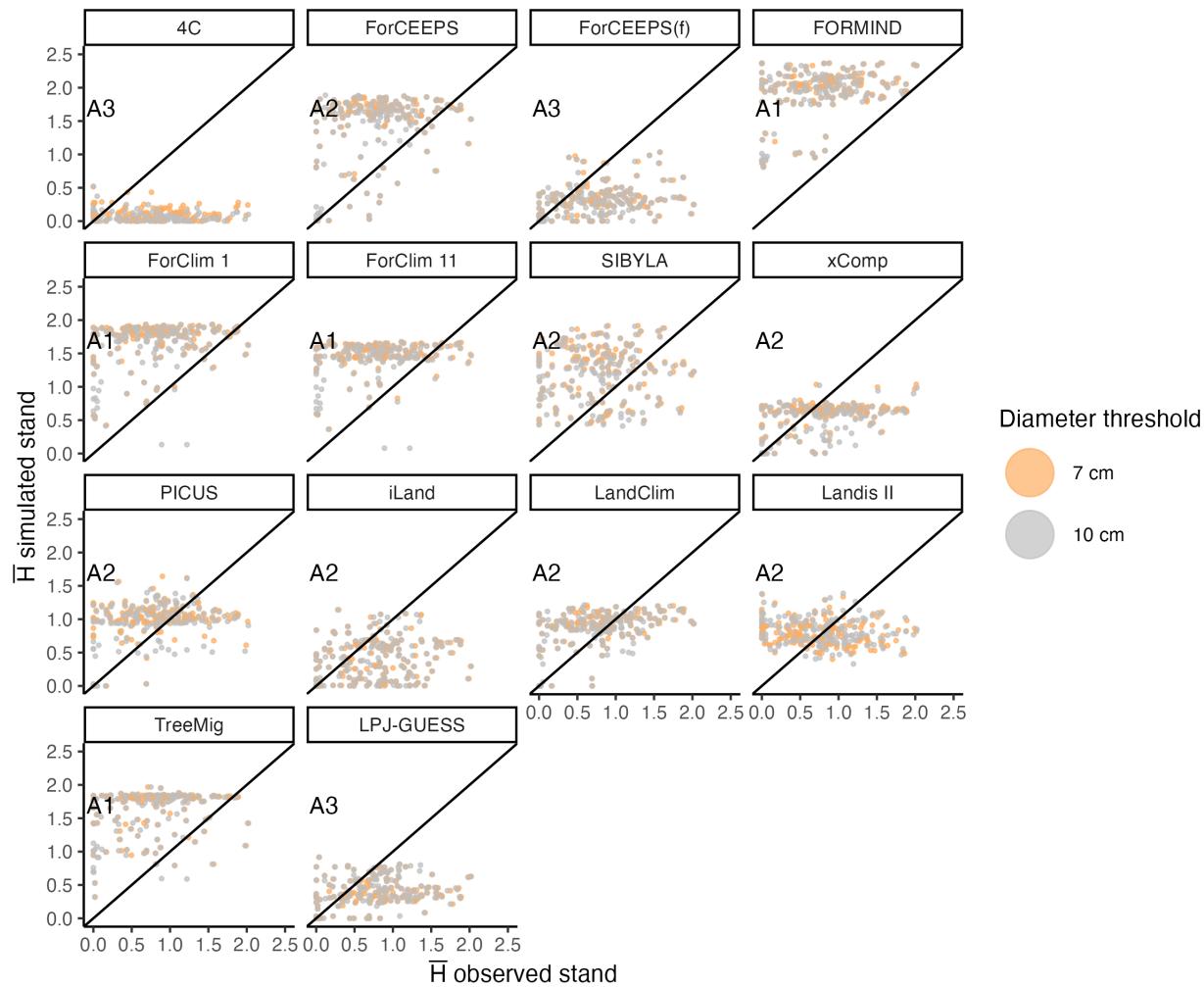


Figure S3: Mean species richness at the stand level per site in the simulated versus observed species richness. The species richness is estimated using the Shannon index calculated by basal area. A1) groups models where the simulated species composition in adults is higher than the observed, A2) groups models where the simulated species composition in adults is similar to the observed, A3) groups models where the simulated species composition in adults is lower than the observed. NOTE: The recruitment threshold recruitment for threshold 7 cm in the empirical data has several sites with no recruitment or no data and the Shannon index for those sites is NA.

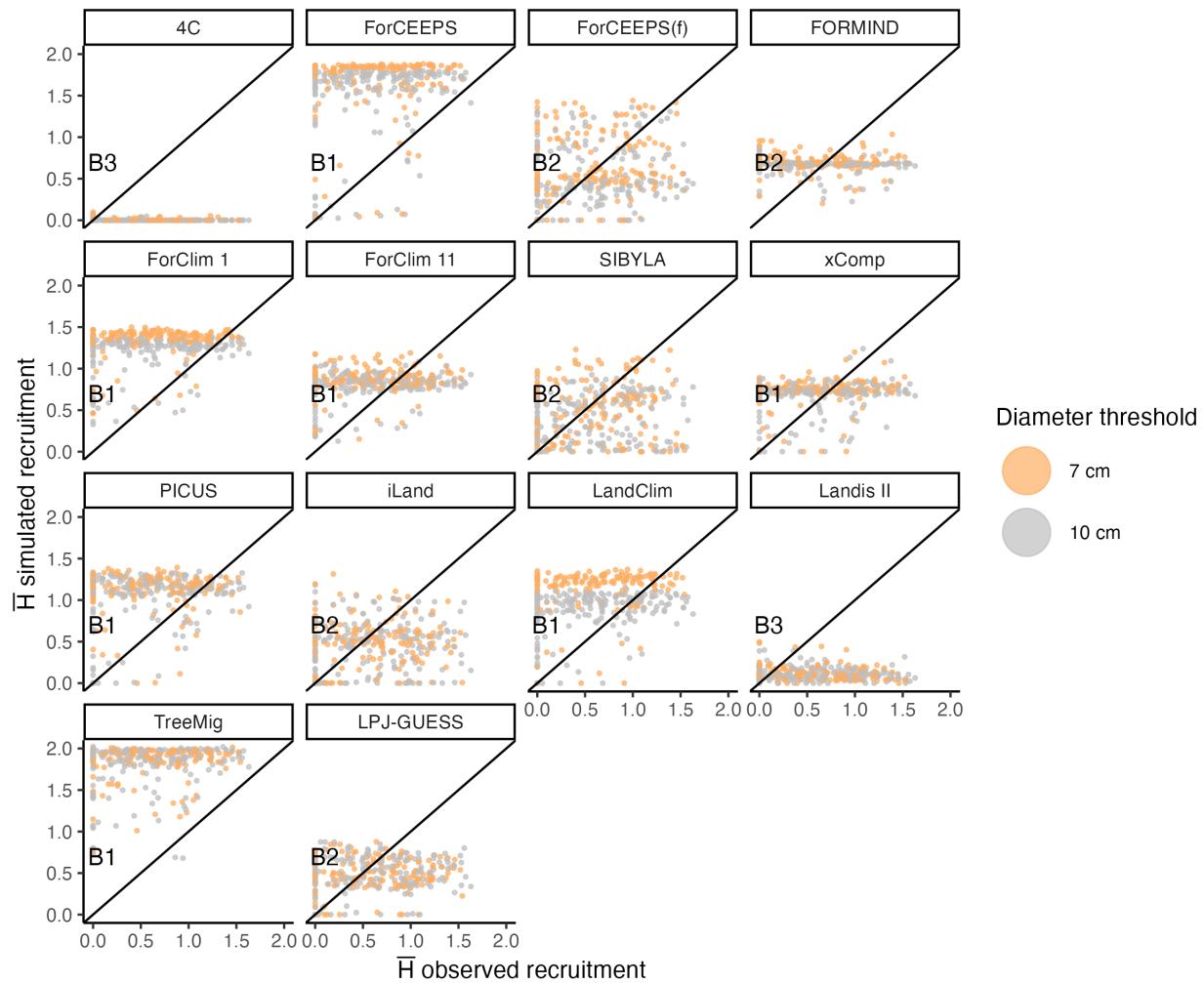


Figure S4: Mean species richness per site in the simulated recruitment versus observed species richness in the recruitment. The species richness is estimated using the Shannon index using the basal area. B1) groups models where the simulated species composition in recruitment is higher than the observed, A2) groups models where the simulated species composition in recruitment is similar to the observed, A3) groups models where the simulated species composition in recruitment is lower than the observed. NOTE: The recruitment threshold recruitment for threshold 7 cm in the empirical data has several sites with no recruitment or no data and the Shannon index for those sites is NA.

Table S3: Test of the significance of the differences of rate of recruitment at 7 cm threshold over 10 cm threshold between the observed values and the simulated values per model.

Y	Group 1	Group 2	p	p.signif
nn710	Observed	4C	0.00291	**
nn710	Observed	ForCEEPS	0.89835	ns
nn710	Observed	ForCEEPS(f)	0.02535	*
nn710	Observed	FORMIND	0.05994	ns
nn710	Observed	ForClim 1	0.59614	ns
nn710	Observed	ForClim 11	0.25164	ns
nn710	Observed	SIBYLA	0.00014	***
nn710	Observed	xComp	1.7e-15	****
nn710	Observed	PICUS	< 2e-16	****
nn710	Observed	iLand	0.07631	ns
nn710	Observed	LandClim	7.7e-05	****
nn710	Observed	Landis II	0.00021	***
nn710	Observed	TreeMig	5.7e-10	****
nn710	Observed	LPJ-GUESS	0.00137	**
nn710	Observed	aDGVM2	0.13992	ns

Table S4: Models trends in mortality between 7 and 10cm (represented in Figure S5), and the ingrowth at 7cm, based in the mean per site. The slope is the rate of change from the linear model between recruitment ratio 7-10 cm and the initial recruitment tree number.

model	Slope	Significance
Observed	0.0038	0.0280
4C	0.0054	0.0145
ForCEEPS	0.0010	0.0009
ForCEEPS(f)	0.0010	0.0000
FORMIND	0.0014	0.4006
ForClim 1	0.0007	0.2244
ForClim 11	-0.0016	0.0000
SIBYLA	0.0537	0.0000
xComp	0.0035	0.0000
PICUS	0.0066	0.0000
iLand	0.0027	0.0000
LandClim	-0.0028	0.0000
Landis II	0.0340	0.0001
TreeMig	0.0006	0.0090
LPJ-GUESS	0.0021	0.0018
aDGVM2	1.6130	0.3680

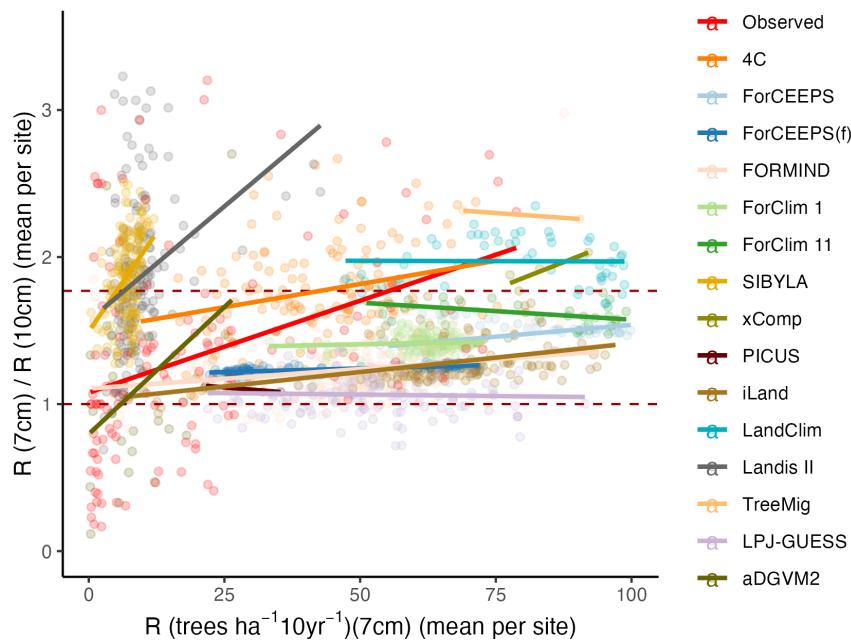


Figure S5: Recruitment ratio between 7 and 10cm and the initial recruitment tree number at 7cm. Dashed red lines mark a ratio equal to 1 indicating no decrease between 7 and 10cm tree recruitment and ratio equal to 1.77 corresponding with the Reineke self-thinning ratio under evenaged conditions.

2.2 Model traits and model performance

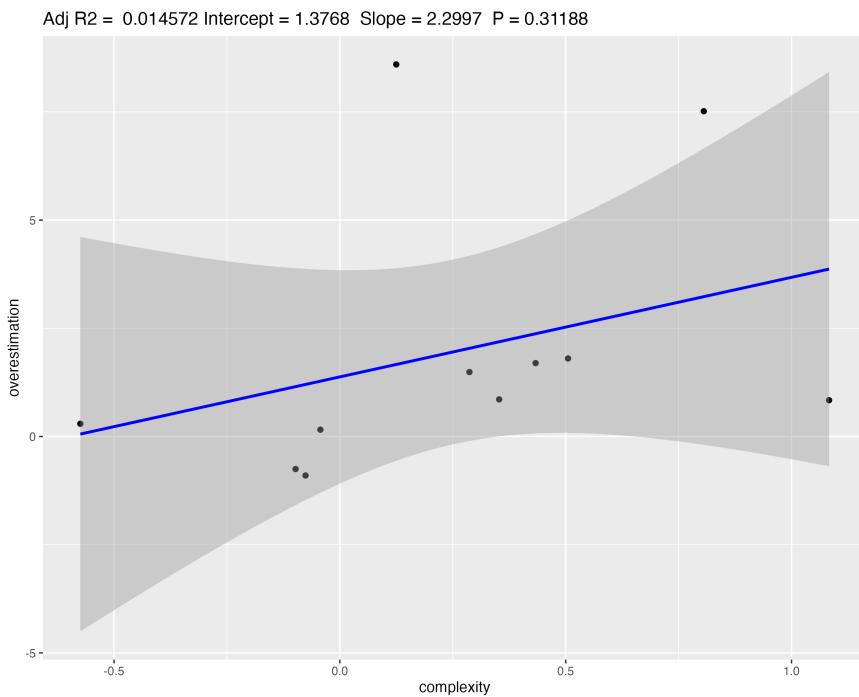


Figure S6: Plot and summary statistics of the linear model between the overestimation proportion in ingrowth levels and the mean complexity value of the establishment modules.

Table S5: Test of the significance of the differences of rate of recruitment at 7 cm threshold of the observed values and the simulated values per model.

Y	Group 1	Group 2	p	p.signif
r.trees	Observed	4C	0.256	ns
r.trees	Observed	ForCEEPS	< 2e-16	****
r.trees	Observed	ForCEEPS(f)	0.027	*
r.trees	Observed	FORMIND	0.036	*
r.trees	Observed	ForClim 1	< 2e-16	****
r.trees	Observed	ForClim 11	< 2e-16	****
r.trees	Observed	SIBYLA	7.9e-09	****
r.trees	Observed	xComp	< 2e-16	****
r.trees	Observed	PICUS	< 2e-16	****
r.trees	Observed	iLand	3.9e-08	****
r.trees	Observed	LandClim	< 2e-16	****
r.trees	Observed	Landis II	7.8e-08	****
r.trees	Observed	TreeMig	< 2e-16	****
r.trees	Observed	LPJ-GUESS	3.3e-07	****
r.trees	Observed	aDGVM2	2.8e-10	****

Table S6: Test of the significance of the differences of rate of recruitment at 7 cm threshold of the observed values and the simulated values per model.

Y	Group 1	Group 2	p	p.signif
ShannonIndexRecruit	Observed	4C	< 2e-16	****
ShannonIndexRecruit	Observed	ForCEEPS	< 2e-16	****
ShannonIndexRecruit	Observed	ForCEEPS(f)	0.17856	ns
ShannonIndexRecruit	Observed	FORMIND	0.07378	ns
ShannonIndexRecruit	Observed	ForClim 1	< 2e-16	****
ShannonIndexRecruit	Observed	ForClim 11	4.5e-10	****
ShannonIndexRecruit	Observed	SIBYLA	0.01610	*
ShannonIndexRecruit	Observed	xComp	0.02709	*
ShannonIndexRecruit	Observed	PICUS	< 2e-16	****
ShannonIndexRecruit	Observed	iLand	0.00462	**
ShannonIndexRecruit	Observed	LandClim	< 2e-16	****
ShannonIndexRecruit	Observed	Landis II	< 2e-16	****
ShannonIndexRecruit	Observed	TreeMig	< 2e-16	****
ShannonIndexRecruit	Observed	LPJ-GUESS	0.00063	***

For Review Only

2.3 Total ingrowth and individual species regeneration niches

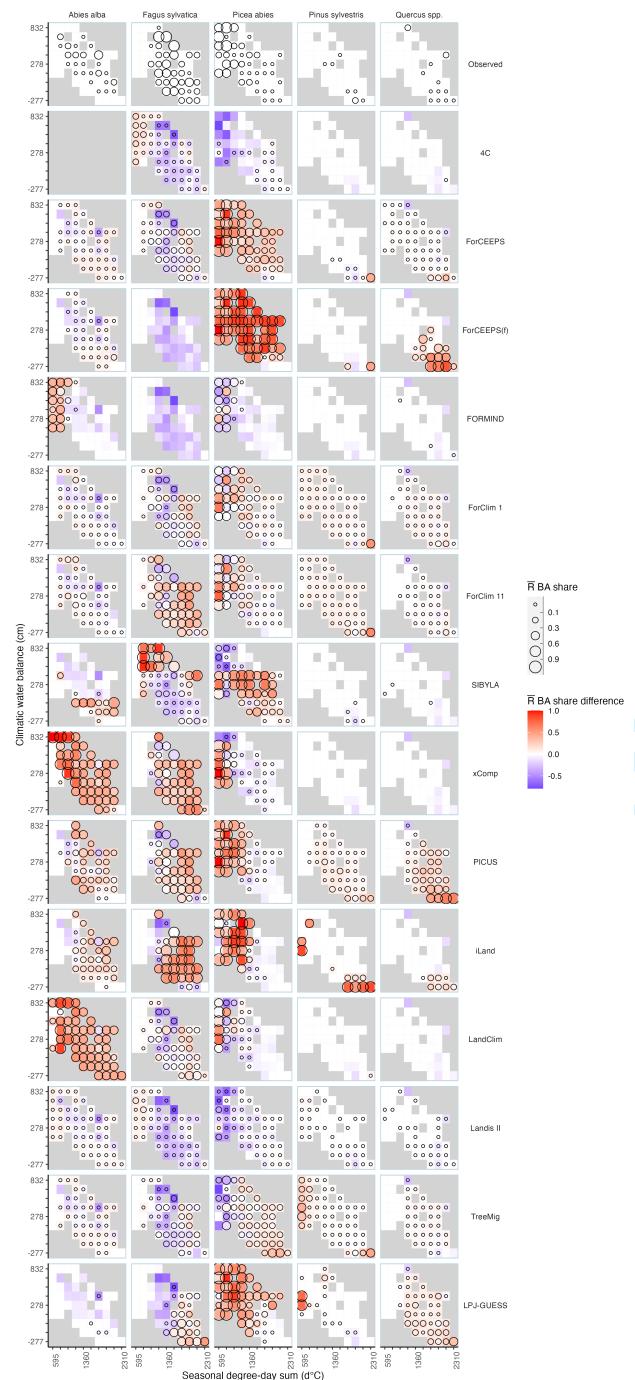


Figure S7: Departure of each model in the representation of the main species recruitment relative abundance from the observations across the environmental gradients for a 10cm threshold. The values shown are the mean of samples per site and across sites in bins, where the range of the climatic water balance (cm) and seasonal degree-day sum ($d^{\circ}\text{C}$) was divided in 10 bins. The point sizes represent ratio between recruited basal area of the species and the total recruited basal area for all the species. The color gradient shows the difference between the simulated and observed ratio between recruited basal area of the species and the total recruited basal area: grey cell represent areas of the environmental gradient in which the model had R BA share and the empirical data did not because both recruitment and stand level recruitments were equal to zero.

Authors Response to Reviewer's Comments

ECOSPHERE Manuscript ID (ECS23-0342) Tree regeneration in models of forest dynamics: a key priority for further research

Olalla Díaz-Yáñez*, Yannek Käber, Tim Anders,
Friedrich Bohn, Kristin H. Braziunas, Josef Brůna,
Rico Fischer, Samuel M. Fischer, Jessica Hetzer,
Thomas Hickler, Christian Hochauer, Manfred J. Lexer,
Heike Lischke, Paola Mairotta, Ján Merganič,
Katarina Merganičová, Tobias Mette, Marco Mina,
Xavier Morin, Mats Nieberg, Werner Rammer,
Christopher P.O. Reyer, Simon Scheiter, Daniel Scherrer,
Harald Bugmann

2023-10-31

Table of contents

1 Response to the Editor	4
1.1 General Comments	4
2 Response to reviewer 1	4
2.1 General Comments	4
2.2 Major comments	5
2.2.1 COMMENT 1	5
2.2.2 COMMENT 2	6
2.2.3 COMMENT 3	7
2.3 Minor comments:	8
2.3.1 COMMENT 4	8
2.3.2 COMMENT 5	8
2.3.3 COMMENT 6	9
2.3.4 COMMENT 7	9
2.3.5 COMMENT 8	10

2.3.6 COMMENT 9	10
2.3.7 COMMENT 10	10
2.3.8 COMMENT 11	10
2.3.9 COMMENT 12	11
2.3.10 COMMENT 13	11
2.3.11 COMMENT 14	11
2.3.12 COMMENT 15	11
2.3.13 COMMENT 16	12
2.3.14 COMMENT 17	12
2.3.15 COMMENT 18	12
2.3.16 COMMENT 19	13
2.3.17 COMMENT 20	13
2.3.18 COMMENT 21	13
2.3.19 COMMENT 22	13
2.3.20 COMMENT 23	14
2.3.21 COMMENT 24	14
2.3.22 COMMENT 25	14
2.3.23 COMMENT 26	15
3 Response to reviewer 2	15
3.1 General Comments	15
3.2 Major comments	15
3.2.1 COMMENT 1	15
3.2.2 COMMENT 2	16
3.2.3 COMMENT 3	16
3.2.4 COMMENT 4	17
3.3 Minor comments:	17
3.3.1 COMMENT 5	17
3.3.2 COMMENT 6	17
3.3.3 COMMENT 7	18
3.3.4 COMMENT 8	18
3.3.5 COMMENT 9	18
3.3.6 COMMENT 10	18
3.3.7 COMMENT 11	18
3.3.8 COMMENT 12	19
3.3.9 COMMENT 13	19
3.3.10 COMMENT 14	19
3.3.11 COMMENT 15	20
3.3.12 COMMENT 16	20
3.3.13 COMMENT 17	20
3.3.14 COMMENT 18	20
3.3.15 COMMENT 19	21
3.3.16 COMMENT 20	21

3.3.17 COMMENT 21	21
3.3.18 COMMENT 22	22
3.3.19 COMMENT 23	22
3.3.20 COMMENT 24	22
3.3.21 COMMENT 25	22
3.3.22 COMMENT 26	23
3.3.23 COMMENT 27	23
3.3.24 COMMENT 28	23
3.3.25 COMMENT 29	23
3.3.26 COMMENT 30	24
3.3.27 COMMENT 31	24
3.3.28 COMMENT 32	24
3.3.29 COMMENT 33	24
3.3.30 COMMENT 34	25
3.3.31 COMMENT 35	25
3.3.32 COMMENT 36	25
3.3.33 COMMENT 37	26
3.3.34 COMMENT 38	26
3.3.35 COMMENT 39	26
3.3.36 COMMENT 40	26
3.3.37 COMMENT 41	27
3.3.38 COMMENT 42	27
3.3.39 COMMENT 43	27
3.3.40 COMMENT 44	28
3.3.41 COMMENT 45	28
3.3.42 COMMENT 46	29
3.3.43 COMMENT 47	29
3.3.44 COMMENT 48	29
3.3.45 COMMENT 49	29
3.3.46 COMMENT 50	30
3.3.47 COMMENT 51	30
3.3.48 COMMENT 52	30
3.3.49 COMMENT 53	31
3.3.50 COMMENT 54	31
3.3.51 COMMENT 55	31
3.3.52 COMMENT 56	32
3.3.53 COMMENT 57	32
3.3.54 COMMENT 58	32
3.3.55 COMMENT 59	32
3.3.56 COMMENT 60	33
3.3.57 COMMENT 61	33
3.3.58 COMMENT 63	33
3.3.59 COMMENT 64	34

1 Response to the Editor

1.1 General Comments

Your revisions should address the specific points made in the review comments. Please note that Reviewer #2 made extensive comments directly on the manuscript.

I should note that personally I can hardly disagree with your basic conclusion that models of forest dynamics need to do a good job of capturing regeneration and recruitment processes. I have devoted most of my research over the past 40 years to developing the field methods to allow incorporation of accurate seedling, sapling and canopy recruitment in the models SORTIE and SORTIE-ND, in systems ranging from the tropics to boreal forests. The specific ecological processes incorporated in those models depended on the goals of the modeling and the ecology of the system, and included processes such as masting cycles, effects of disturbance on seedbed substrate dynamics, seed and seedling predation by small mammals, browsing by ungulates, Janzen-Connell effects, and, of course, detailed analyses of shading and neighborhood competition. More recently, this includes developing statistical models of seedling recruitment from the US forest inventory network that have been directly incorporated in the model for forests of the entire eastern US. To my mind, the difficulties the models you reviewed encountered in capturing canopy recruitment speak more to the limitations of the approaches used in model development than in any inherent difficulties in capturing regeneration and recruitment in models of forest dynamics.

Response to editors comments

Thank you for taking the time to review our manuscript and for your comments. Our main take-away from this manuscript is that while it would certainly be possible to calibrate all 15 models against the EuFoRJa data, it is important to ensure that this aligns with the study's goals. If the goal is to provide extrapolation into a future climate, calibrating against data under the current climate may not be reliable, as shown by the two models in our set that *were* calibrated against such data (i.e., SIBYLA and xComp). This is why we believe it is valuable to test the performance of the models "as they are", i.e. as we have done in this study.

We have carefully reviewed and addressed each of the specific points made in the review comments in this revised version, as explained below.

2 Response to reviewer 1

2.1 General Comments

The manuscript compares how regeneration and recruitment is simulated in 15 forest models, using a large network of forest reserves in Europe as the observed data. I want to start by saying there are many strengths of the paper. I was impressed with the comparison of 15

models, that the observed data was not used to parameterize or fine-tune the models, and the high quality of the writing. The introduction and discussion were easy and enjoyable to read, and the topic is critical and relevant. Some of the details about the models and methods were placed in the very substantial appendix, and I agree with their choices. However, the manuscript is still quite long with lots of multi-paneled figures. If the journal does not have a page limit, then this is not an issue. If there is a page limit, then some tough choices will have to be made for which figures to keep. Figures 2, 3, 6 and 7 are the most important (in my opinion).

I have a few major comments. One of which requires a different statistical analysis – which is why I do not consider this a ‘minor revision’. However, there are no fatal flaws, and all of my comments are fixable. I hope they serve to improve the manuscript.

Our general comments to reviewer 1:

Thank you for taking the time to review our manuscript. We have made changes to the manuscript based on your feedback and believe that these changes have improved the overall quality of our manuscript.

We appreciate the concern about the length of the manuscript and the number of figures included. We have made an effort to reduce the text where possible and have verified that we are following the journal guidelines. Despite our best efforts, however, it has been difficult to significantly reduce the length of the manuscript.

Our responses to each of the comments and the concrete actions we have taken are detailed below.

2.2 Major comments

2.2.1 COMMENT 1

‘regeneration’ versus ‘recruitment’. I personally have had this conversation multiple times over the years, so I understand that there is variation in the definitions and uses. However, the authors do define these terms in the methods (L167-171) and I completely agree with them. Regeneration is for flowers, seeds, and seedlings while recruitment is for reaching a certain size threshold. However, the entire manuscript refers to regeneration as “passing of trees across a specific diameter threshold (“ingrowth”) (L142), either reaching a size of 7 or 10 cm DBH. This is recruitment. Since this is the response variable for the manuscript, I would suggest replacing ‘regeneration’ with ‘recruitment’ throughout the manuscript.

Response to comment 1:

Thank you for pointing this out. We discussed terminology several times and basically agree. We have followed your suggestion and have replaced the erroneous use of ‘regeneration’ by ‘recruitment’ where we refer to ingrowth into a certain DBH class. However, in some places

the term ‘regeneration’ is more accurate, especially in many parts of the introduction and discussion sections. ‘Regeneration’ is a broader term that starts from flowering and ends with ingrowth into a given size class; that is, recruitment is the result of successful regeneration. Thus, where we refer to the entire process of regeneration, it would not be appropriate to use ‘recruitment’, and therefore we maintain the former term where it is fitting.

2.2.2 COMMENT 2

The difference between 7 cm and 10 cm DBH in the observed data. From the methods, there is one dataset that has sites where recruitment was defined as 7 cm DBH, and then another where recruitment was somewhere between 7 – 10 cm DBH. It was not clear how these were treated and compared in the results. I assume that the models simulated all 200 sites, but then the recruitment results were compared like-to-like. So, if a site defined recruitment as DBH = 8.0 cm, then all 15 models would compare simulated recruitment at that site only at 8 cm. Since the data are boxplots, I am not able to tell. For example, Figure 2 – there should be 165 points for all of the 7 cm bars (observed and simulated) and 35 data points for the 10 cm bars (observed and simulated). If true, then please clarify. If not true, then please explain. Also, how was the data analysis performed when looking at the change in mortality and species diversity between 7 to 10 cm? I assume only the 165 stands were used. Please clarify.

Response to comment 2:

We realized that there is an issue with the way we communicated our diameter thresholds approach in the manuscript. The observed data had varying diameter thresholds for reporting tree recruitment, and, for the purposes of this study, we merged the recruited trees values into two thresholds of 7 and 10 cm. That is, in all plots ($n = 165$) that had a caliper limit <7 cm, only the successful recruitment events at 7 cm were considered; and analogously for plots ($n = 35$) with a calliper limit between 7 and 10 cm, where only recruitment at 10 cm was considered. Thus, the 7 cm dataset ($n = 165$) actually is a subset of the 10 cm dataset ($n = 200$), as all data from the 7-cm threshold could be re-calculated for the 10-cm threshold. Thus, there was no ‘like-to-like’ comparison for any other diameter threshold, only for 7 and 10 cm.

When examining the change in mortality between 7 and 10 cm, or the difference in species diversity between observed and simulated data (Figure 4), we could only use the 165 sites that contained both pieces of information at the 7 and 10 cm thresholds in the observed data. This is reported in the supplementary materials software, script “‘Figures.R““ line 536 and L284-349 and now better explained in the manuscript: we have better explained the difference on the observed data in the section 3.2 Observed data L[225-230] and we have added a new text in the caption of figure 4 L[1085-1086].

2.2.3 COMMENT 3

The t-tests. The results in Table S2 and S3 (comparing at 7 cm and 10 cm DBH) should be at least by paired t-tests not two-sided t-tests (L319), where you are comparing the difference within the same stand at 7 and 10 cm. The results Table S5 and S6 should not be a t-test at all. Performing 15 individual t-tests is a classic Type-1 error situation (technically this applies to Tables S2 and S3 as well, a comment on that below). For Tables S5 and S6, I would suggest a general linear mixed effects model. You can take the difference between the observed, and the simulated output and make this the response variable (Y). Then, include model complexity, feedback, model type and scale as fixed explanatory variables. And include the stand ID as the random effect. This will more accurately answer your question about if any of these fixed explanatory variables can predict over or under representation of recruitment. For Tables S2 and S3, you could do the same thing (i.e., calculate the difference between the observed and the simulated output for each model) and make this your response variable. But the only fixed effect would be the actual model, and then you could do an ANOVA (again, stand ID as the random effect) and then a post-hoc test to see which models were or were not different from observed. If you do want to keep the 15 paired t-tests, then you need to account for Type I error rates and adjust your p-values accordingly.

Response to comment 3:

The objective of running the t-tests was to allow us to discuss differences between mean values per site between models, and observations and simulations between the 7 and 10 cm diameter thresholds. The t-test, as employed in our study, is not intended to serve as 'proof' of our findings, and we have carefully avoided an over-interpretation of their outcomes. Rather, they serve as a practical tool for handling the extensive number of comparisons in our study. To clarify this, we have included a statement in the manuscript, cf. L[336-337].

Following the reviewer's comment, we have included a correction for multiple tests, although we are aware that all corrections have pros and cons. We used the Bonferroni correction to make sure that the actual Type I error probability for the combined H0 is smaller than or equal to the nominal test level.

We have also looked into the alternative approaches suggested by the reviewer. We cannot do a paired t-test as for paired samples it is necessary to have the same number of subjects in both groups (observed and simulated), which is not the case here, as for example some models did not simulate all the sites. Yet, we tested the possibility of creating a GLME using as random effect the site and as explanatory variables all the features that we were interested in (model complexity, feedback, model type and scale), as suggested by the reviewer. However, this approach does not provide the comparison of observed vs. simulated data that we are interested in. Following the reviewer's comments, we conducted further tests to see if we could better explain differences between observed and simulated values using other approaches. We tested a mixed model to account for the randomness induced by to site (plot) differences. For example, a

GLME with recruitment values as the response variable and models as the explanatory variable, with observed data as the reference (with a Poisson distribution because we are dealing with count data). This worked well, and the patterns were the same as in the t-test (except for the model 4C). We also considered pursuing the same approach with the Shannon index and the differences in mortality, but in this case we think it is more relevant to understand the mean differences across sites, and therefore a t-test is more suitable. We furthermore tested a linear mixed model to assess differences between the 7 and 10 cm thresholds regarding the Shannon index, by creating a new parameter where the explanatory variable was a combination of the model and the diameter threshold. We then ran pairwise multiple comparisons to determine which means are different. Also here, the patterns were quite similar to the ones resulting from the t-test except for two landscape-level models (TreeMig and Landis-II) that were significant in this test, contrary to the t-test.

After considering all the options that the reviewer suggested and testing others, we have made some adjustments to the t-test analysis, with the correction for multiple testing and cleaner output tables, as suggested by the reviewer. In addition, we have explained the purpose and interpretation of these tests better on lines [336-337].

2.3 Minor comments:

2.3.1 COMMENT 4

The abstract is not clear if the models do a good or poor job in simulating regeneration (L59, that they match observed ranges, and then L66, considerable mismatch between simulations and observed data, but that they do capture tree regeneration). Please clarify the main message in the abstract.

Response to comment 4:

We have changed the abstract based on feedback from both reviewers, including clarifying certain parts for better understanding. The models were able to capture some aspects of regeneration more effectively than others. One example, is the sentence highlighted by the reviewer that was referring to the diversity of regeneration.

2.3.2 COMMENT 5

L123-138: this paragraph has a lot of “on the other hand” type of statements, and it does not lead nicely into the next paragraph about the current study. It was a bit of a surprise (i.e., I did not expect the main purpose of the study when I started to read the following paragraph). It seems like the main issue is that we just don’t know how uncertain regeneration and recruitment methods are, because we lack the comparison across models and a comprehensive empirical recruitment data set that we can use to compare. And that is what your study is doing.

Response to comment 5:

We have made several changes in this paragraph to improve its readability and flow of ideas towards explaining the main purpose of the study.

2.3.3 COMMENT 6

Table 1. Please clarify what the letters are for (a, b, c and d in the table headers). For models that start from saplings or ingrowth, please add in the size as well (i.e., start from saplings with DBH = 1.0). Could also add in the classification scheme mentioned in the methods (L168, regeneration or recruitment) to Table 1.

Response to comment 6:

We have added a new column classifying the models into recruitment or regeneration models; and the letters a, b, c, d, e and f in the table are now better described. Regarding the suggestion of adding the size of the saplings or ingrowth: we are concerned that this would be more confusing than helpful. For example, there are models that start from saplings without a diameter threshold, as they are biomass-based (e.g. LandClim), and providing all information for each model would make this table rather overloaded, without a clear benefit, as we believe. Therefore we have decided not to include this information. We would also like to note that the 15 model reports in the Supplementary Materials provide more details how regeneration is simulated.

2.3.4 COMMENT 7

L218 – do I understand this correctly, that regeneration rates are a mean of 56 trees per ha per decade? This seems really low (assuming that this is for >10 cm DBH).

Response to comment 7:

Yes, this is the mean value across all observations and sites. There are two major reasons why this number appears low. First due to the zero-inflated nature of the data, and second because these are forest reserves, typically being characterized by denser forests in which regeneration can be expected to be low. This value is in agreement with further observed data from reserves (Käber, Y., et al (2023). Sheltered or suppressed? Tree regeneration in unmanaged European forests. Journal of Ecology, <https://doi.org/10.1111/1365-2745.14181>).

2.3.5 COMMENT 8

L264 – 274: how will this be accounted for later, that not all models simulated the same set of species?

Response to comment 8:

For those models that simulated more than the eleven species, we created a different category “others” where we lumped those extra recruited species. These were counted when we calculated the diversity index as well as for total recruitment values. We have added this information in the simulation protocol section L[278-281].

2.3.6 COMMENT 9

L340: “regeneration basal area of that species” is mentioned twice, please delete the second one.

Response to comment 9:

Following the reviewers comment we have modified the text L[354-357].

2.3.7 COMMENT 10

L364: so please remove the 4C model from the figure as it is not comparable. Do you need to also remove other models that don't simulate the same number of species too?

Response to comment 10:

We acknowledge the point that the reviewer is making, but even though it is not directly comparable, we still believe to be insightful for the reader that we share the findings related to the 4C model. Note that in those models that simulate more species, these additional species were lumped in the category ‘others’ (as explained further above), thus with little impact on the comparison.

2.3.8 COMMENT 11

Table S2 and S3: What are the 3 columns for p-values? Please report just the one you used. No need to have a column to state ‘t-test’ – just move this to the table caption. Then add in a column for df, the test statistic (t) and the effect size (was it higher or lower and by how much?). NOTE, this comment may become irrelevant if you change the statistical analyses.

Response to comment 11:

We have cleaned all the output tables.

2.3.9 COMMENT 12

Figure S1 and S2 – I can't tell the difference between observed and simulated. Are they different colours? The legend should be larger. What is the black line for?

Response to comment 12:

We have increased the font size of the legend, and now the difference between observed and simulated is visible. We have also explained the meaning of the black line in the caption.

2.3.10 COMMENT 13

Figure 4 (and Figure S3) are interesting in that there are no patterns with overstory diversity, and that this was not explained by feedback with seed production. What happens if you colour the points, by those from models with the feedback? Do you address in the discussion why this feedback makes no difference to recruitment?

Response to comment 13:

We agree that it is interesting that the models with feedback did not capture better the species diversity of the regeneration compared with those without, with the exception of ForCEEPS. Figure 4 shows three representative patterns observed across models using 6 of the models as an example, and Figure S3 provides the details for each model. We believe it is not very informative to color the points from the models with feedback in Figure 4 as this will only color all the points from some of the panels and it is possible to identify the models with feedback in table 1. We have addressed this point in the discussion section L[563-570].

2.3.11 COMMENT 14

All figures in the Appendix, please make the X and Y axis labels larger and centered. I actually thought labels were missing initially and it took a while to find them.

Response to comment 14:

We have corrected these figures.

2.3.12 COMMENT 15

L400 – 402: where is this conclusion about mixed species or monocultures coming from? I don't see that in the figures or the results.

Response to comment 15:

This paragraph is addressing the comparison with the theoretical Reineke value which is usually calculated for even-aged, single species stands. We have change this text to clarify this aspect L[324-330].

2.3.13 COMMENT 16

Figure 6 is excellent.

Response to comment 16:

Thank you!

2.3.14 COMMENT 17

Table S4: please clean up. No need to report 7 places after the decimal point. Please clarify in the table caption what the slope represents (maybe mention Figure S5 too, as I assume it is illustrating the slopes). Another suggestion, would be to make the slopes solid lines if significant, or dashed lines if not significant. Not sure if that would help, but if it does then it could eliminate the need for this table.

Response to comment 17:

We have cleaned this table and changed the table caption including a reference to Figure S5 as suggested by the reviewer.

2.3.15 COMMENT 18

Not sure both Figure 7 and Figure S7 are necessary (I admit I spent a few minutes going back and forth between them trying to understand how they differed). I would just keep Figure 7 and note that the scales differ.

Response to comment 18:

We have removed figure S7.

2.3.16 COMMENT 19

L547-548: predicting stand diversity, did I miss this? I don't believe Figure 4 shows this, as if you put all data points on the same panel it would just be a random scatter of points that covers the entire space (not necessarily closer to the 1:1 line).

Response to comment 19:

Figure 4 shows three examples of the patterns observed across models (overestimation intermediate and underestimation) using one exemplary model of each category. The actual patterns per model can be found in Figure S3. In the results, we showed that most models capture reasonably well or overpredicted species diversity at the stand level, as most of the models are following the patterns presented in plot 1 and 2 (part A of Figure 4). We have decided to keep the referenced statement regarding this finding in the discussion.

2.3.17 COMMENT 20

L575: that most models did not deviate exceedingly from observations seems in contrast to what the previous section was stating. What is this based on?

Response to comment 20:

In this statement we were referring only to regeneration levels. Although we agree with the reviewer that we did state that most of the models overestimate regeneration, here our point is that even though they did overestimate regeneration, they did not deviate exceedingly, especially given that they have never been confronted with this type of data.

2.3.18 COMMENT 21

L584: What does this mean, about degrees of freedom for modelling regeneration?

Response to comment 21:

With this statement, we were referring to the multiple potential combinations of approaches of modeling tree regeneration. We have modified this for clarity L[596-598]

2.3.19 COMMENT 22

L624: Missing the end of the sentence.

Response to comment 22:

This was a glitch, which now is corrected L[636-638]

2.3.20 COMMENT 23

L647: Completely agree that observed data is a snapshot of a stochastic and dynamic system and care should always be taken with comparing to simulation data, however I disagree about the following stamen (L648). The models should be able to capture broader patterns along climatic or environmental gradients.

Response to comment 23:

We have removed these statements.

2.3.21 COMMENT 24

L655: But the landscape models generally did about the same (i.e., not consistently worse) than the rest, correct?

Response to comment 24:

Yes, the landscape level models were not performing significantly worse than the others. However, they were not performing better either, potentially because they were used in “single site” rather than “landscape” mode. Thus, the explicit consideration of horizontal space would possibly allow the landscape level models to better represent regeneration processes, as they are designed for that. This, however, was completely out of the scope of the present analysis, as there are no landscape-level data of the surroundings of the 200 plots.

2.3.22 COMMENT 25

L663: But in general, the models were more sensitive to water balance and the observed data did not have any real pattern. Why would this coarser representation of water balance in the model, cause it to be more sensitive?

Response to comment 25:

Our point here is that the models, although many of them have detailed schemes for the water balance, may not have performed well because of the absence of site-specific (truly reliable) soil information.

2.3.23 COMMENT 26

L675: what would a 'comprehensive regeneration dataset" look like? There are large data sets for seed production (MASTIF, MASTREE, etc.) and seedling data from national forest inventory programs. Does EuFoRIa (L677) have everything you need? Is it just a matter of expanding the spatial and temporal extent? Please explain what sort of data that the models would need (imagine an ideal world).

Response to comment 26:

We have changed this text L[684-688]. Although there are multiple data sources (such as MASTIF, MASTREE, National Forest Inventories, Forest Reserve plots, etc.), usually these data do not refer to the same locations and thus cannot be used in a site-specific manner for calibrating or testing models that intend to follow the "regeneration" chain of processes.

3 Response to reviewer 2

3.1 General Comments

It is valuable to demonstrate that leading forest dynamics models are unable to characterise regeneration processes very well. I also thought the "blind trial" approach was clever. Huge effort must have gone into running 15 models under various starting conditions and synthesising the outputs into a coherent paper. Choices have been made about which details to include in the main body of the text vs supplementary information and I thought a good balance was struck, although I have made several suggestions on the pdf.

Our general comments to reviewer 2:

We would like to thank Reviewer 2 for this positive overall assessment and the level of detail of the review. We have taken care to go through all suggestions made in the PDF, and have changed/improved the manuscript in most instances. Our responses to each of the comments and the concrete actions we have taken are detailed below.

3.2 Major comments

3.2.1 COMMENT 1

Where the paper falls short, in my opinion, is in providing a clear path forward. It shows that the models are too simplistic to simulate regeneration processes accurately but provides few insights into how we can improve their performance. The paper reports that creating more complex models is not valuable in and of itself, but that's not to say that intelligent improvement of specific subcomponents may hold the key. For instance, deer browsing and

competition with herb layer plants may well contribute to lower recruitment being observed than simulated, but how can we demonstrate that, and if the inadequate representation of these processes in the models is the problem, what should be done to improve the models? The review by Hanbury-Brown et al. (2022) on “future forests within Earth system models: regeneration processes critical to prediction” provides a good summary of current knowledge and future directions, including thoughts on 1. reproductive allocation and seed production; 2. dispersal; 3. seed survival, germination, and resprouting; and 4. seedling survival and growth. I would encourage the authors to frame their discussion in a similar way. Advanced statistical approaches that bring together simulation models and Approximate Bayesian Computation to estimate parameters of recruitment submodels may hold the key (e.g. <https://onlinelibrary.wiley.com/doi/full/10.1111/ecog.04824>). To me, a key question is how we can make better use of field measurements and statistical analyses to refine forest dynamics models, i.e. not only parameterise them but also to identify which submodels need refinement and the function that need inclusion.

Response to comment 1:

In this first assessment, we were not aiming to identify what specific subcomponents may hold the key. Therefore, based on the current approach, we could not provide more concrete recommendations for model development, which would most likely be specific for each of the 15 models. Maybe we need more such studies, including calibration as a first step, and then application against a withheld part of the data or application (and comparison) under climate change scenarios to evaluate the robustness of the signal under future conditions. We agree with the reviewer that we need more insights into how we can better simulate regeneration processes in each of these models, but this was beyond the scope of this paper.

3.2.2 COMMENT 2

I felt the readability of the paper would be improved by introducing a “road mapping” paragraph at the end of the introduction or start of the methods section that sets out the approach in broad terms.

Response to comment 2:

We have added a “road map” at the end of the introduction L[159-163].

3.2.3 COMMENT 3

The words “ingrowth”, “regeneration” and “recruitment” are used interchangeably to describe the recruitment of new trees into the 7-cm diameter class. I would encourage the authors to use the term “ingrowth” or “recruitment” to this to describe this process, and keep “regeneration” as an overarching term describing the entire process of seed production, dispersal, early establishment, and onward growth.

Response to comment 3:

Thank you for pointing this out. We discussed terminology several times and basically agree. We have followed your suggestion and have now use recruitment to discuss the process and regeneration as the overarching term.

3.2.4 COMMENT 4

The article is written entirely by Europeans, unless I'm mistaken, and very much focused on the central/eastern European forestry literature. I would encourage a deeper dive into the relevant North American and Asian literature when introducing and discussing their work.

Response to comment 4:

Indeed, most of the people involved in this manuscript have a central European focus in their current work, and so did the study setup. We have included global references or references with a focus outside Europe when they were relevant. We have evaluated whether this could be further “internationalized”, but felt our choices were appropriate.

3.3 Minor comments:

These comments have been extracted from the reviewer 2 comments on the PDF document of the manuscript.

3.3.1 COMMENT 5

L51-52 “However, an assessment of their ability to accurately represent tree regeneration is lacking.” Not true.

Response to comment 5:

We have modified this sentence L[52-53]

3.3.2 COMMENT 6

L56 “The results are evaluated against comprehensive data from unmanaged forests.” comprehensive?

Response to comment 6:

We have changed the word “comprehensive” to “extensive” to express that we used a wide-range data that considers many important elements of regeneration L[57]

3.3.3 COMMENT 7

L56 “Models simulating higher species diversity at the stand level do not feature higher regeneration diversity.” feature?

Response to comment 7:

We have kept “feature” in this case.

3.3.4 COMMENT 8

L76-78 “A wide range of models of forest dynamics were developed over the past decades considering the impacts of climate”

Response to comment 8:

We have changed the word “were” with “have been” L[78].

3.3.5 COMMENT 9

L79 K. Vanclay and Skovsgaard 1997;

Response to comment 9:

We have corrected the reference style L[80].

3.3.6 COMMENT 10

L83-84 “Which is a clear research gap in the context of climate-induced forest disturbances and forest resilience.” incomplete sentence

Response to comment 10:

We have changed this sentence L[84-85].

3.3.7 COMMENT 11

L85-86 “Tree regeneration arises from multiple processes such as pollination, fruit maturation, seed production, dispersal, germination, juvenile growth and survival” including

Response to comment 11:

We have changed “such as” with “including” L[86].

3.3.8 COMMENT 12

L90-92 “Currently, tree regeneration processes in dynamic forest models are handled in a multitude of ways (König et al. 2022; Bugmann and Seidl 2022): from 1) entirely ignoring it (as done in classical forest growth models, e.g., Pretzsch et al. 2002),” it.

Response to comment 12:

We have changed “it” with “them” L[92].

3.3.9 COMMENT 13

L96-98 Particularly given your conclusion that the models are poor at predicting regeneration processes, should we now return to statistical parameterisation afresh and find better ways of extrapolation/interpolation across space/time?

Response to comment 13:

It is indeed interesting to account with the increasing data availability to better understand tree regeneration processes. However, we believe that the main take away it is not that it would be impossible to calibrate all 15 models against e.g. the EuFoRIa data to better capture regeneration, but that the focus should be in the study goals. And when the goal is to provide extrapolation into a future climate, then calibration against data under current climate may be unreliable. With the approach we have followed in this study where we have asked the models to simulate forest dynamics and capture regeneration without statistical parametrisation we can better assess their behavior and understand how their current model structures and inclusion of exclusion of certain ecological processes impact the outputs we are seeing.

3.3.10 COMMENT 14

L100-101 “Overall, models are needed to...(3) identify the most important processes that are shaping ecological patterns.” doesn’t that depend on what goes into the model?

Response to comment 14:

We are not completely sure of the reviewers point here. If the reviewer is referring to the fact that the most important processes will be those included in the model, it is true that you can only asses those processes included or how including or not another processes impact regeneration. This was also the idea with considering a large pool of models, that we could assess how their structural complexity, or the inclusion/exclusion of a certain process affects the results we are seeing. This also goes very much in line with our research recommendations presented in the discussion where we proposed that improvement of the regeneration modules is implemented as additional features that can be traced back, as done here for the variants of

ForClim and ForCEEPS, and that model complexity and structure must always be connected with modeling objectives.

3.3.11 COMMENT 15

L101-106 “Given the current strategies that are used in models of forest dynamics to represent tree regeneration, model behavior often is prone to problems, such as very high levels of tree regeneration that necessitate excess mortality at early stages of tree life to simulate correct stand structure and composition.” ref? tens of thousands of seeds...

Response to comment 15:

We have added a reference here.

3.3.12 COMMENT 16

L112-114 “A related issue is the excessive reduction of species diversity due to positive feedback effects, such that eventually just single-species stands remain.” ref

Response to comment 16:

We have added a reference here L[105].

3.3.13 COMMENT 17

L127 I suggest making better use of the hanbury-brown 2022 review.

Response to comment 17:

We currently use the exciting Hanbury-brown 2022 review in several places across our manuscripts, following the reviewer’s comment here and the general comment 2 we have also get inspired by their structure by creating a road map to help the reader to navigate the manuscript content.

3.3.14 COMMENT 18

L145-147 “Due to the large variability in tree regeneration patterns in nature and the large number of factors driving this process including some that are not incorporated explicitly in most models, such as deer browsing—”

Response to comment 18:

We believe the reviewer highlighted the words “models, such as deer browsing” because they think this is not correct. From our sample pool only 4 models incorporated explicitly deer browsing.

3.3.15 COMMENT 19

L185 “(mean regeneration formulation complexity across all processes >0, Table 1).” Meaning

Response to comment 19:

This statement relates to Bugmann and Seidl (2022) where complexity values were provided. The exact details of how we calculated this mean regeneration formulation complexity value are available in the software supplementary materials to this paper: “‘figure.R’“ line 706-750. This value represents the mean complexity of regeneration formulations per model.

3.3.16 COMMENT 20

L186-188 “ForClim variant 1 (Bugmann et al. 1996) is based on a recruitment module that adheres closely to the concept introduced by Botkin, Janak and Wallis (1972)” which is...

Response to comment 22:

We have added information to briefly explain the general approach L[196-197].

3.3.17 COMMENT 21

L197-198 “Regeneration data covering a wide range of environmental conditions are hard to obtain, and this is one of the reasons why most models of forest dynamics have never been confronted with a dataset” not sure that’s true. Some of the large forest inventories include small seedlings and saplings in nested plots, which are a step in the right direction.

Response to comment 21:

We agree with the reviewer that permanent plots such as those present in National Forest Inventories could potentially be a good source of information to obtain regeneration data. However these often are collected from managed sites, do not capture ingrowth below 10cm, and are hard to harmonize between different inventory strategies.

3.3.18 COMMENT 22

L200-202 “The observations used here are derived from a novel and unprecedented network of sites in forest reserves that represent the range of environmental gradients in temperature and precipitation in Central Europe as compiled in the framework of the EuFoRIa network (EuFoRIa 2019)” I don’t feel “novel and unprecedented” adds anything to this sentence. There are plenty of networks outside Europe providing this sort of data

Response to comment 22:

We agree with the reviewer’s comment that there are other forest inventory data networks outside Central Europe even on forest reserves that could provide this sort of data. However, within Europe no such data existed before. The research network (EuFoRIa) was established in 2019 and therefore is in fact novel, and within Europe such data is also without precedence.

Yet we acknowledge that our statement needs to be put in a European context much clearer. We hope that our adjustments to the text resolve the exaggeration of our initial formulation L[211-213].

3.3.19 COMMENT 23

L208-209 “We selected 200 sites from this network as the benchmarking dataset for the simulation to be representative of the environmental variation contained in the data.” COOL

Response to comment 23:

We agree :)

3.3.20 COMMENT 24

L211-212 “exposition (i.e.,slope and aspect).” unfamiliar term to me

Response to comment 24:

We have changed “and exposition (i.e.,slope and aspect).” with “, slope and aspect.” L[223].

3.3.21 COMMENT 25

L212-213 “Regeneration thresholds for these sites differed between diameters of 0 and 10 cm” minimum size threshold? and stem diameters

Response to comment 25:

Following the reviewer’s comment we have changed this sentence L[224-225].

3.3.22 COMMENT 26

L215 “another 35 sites with diameter thresholds between 7 and 10 cm.” OK, so we’re talking about v large trees here

L218 “featured 30,900 newly established trees” so they are trees....

Response to comment 26:

The minimum threshold that could be standarized across reserves was of at least 7cm.

3.3.23 COMMENT 27

L215 “on this unique dataset” not so unique

Response to comment 27:

We have changed this text L[235].

3.3.24 COMMENT 28

L221 “nformation, cf. Käber et al. (2023).” in review and not available

Response to comment 28:

This paper have recently being published and the reference has been updated accordingly.

3.3.25 COMMENT 29

L223 “The overarching goal of the experiments was to assess the tree regeneration as it arises” simulation experiments and remove “the”

Response to comment 29:

Changed as suggested L[238-239].

3.3.26 COMMENT 30

L225 “tree regeneration as the passing of a breast height diameter threshold of 7 or 10 cm, respectively” I suggest “tree recruitment” not “tree regeneration” here. respectively??

Response to comment 30:

We have removed “respectively” L[240].

We have re-assessed the terms used across this manuscript and change it across the text accordingly.

3.3.27 COMMENT 31

L 227-228 “simulations, providing input variables” provided with

Response to comment 31:

We have removed “providing” L[242].

3.3.28 COMMENT 32

L229 “Neither were further site information (except for the data specified below) nor any data” These were “blind trials”: modelling group were not provided with site information ...

Response to comment 32:

The reviewer is correct that we did not provide e.g. site coordinates (with the exception of iLand and aDGVM2 that needed this for a model specific soil input, this is described in the manuscript), but we did provide topography, climate and soil site specific variables.

3.3.29 COMMENT 33

L230-231 “That is, the models were run in”blind flight” mode.” Remove

Response to comment 33:

Changed as suggested.

3.3.30 COMMENT 34

L237-238 “Soil quality data were provided as continuous values between 1 and 5” ?hydrological quality?

Response to comment 34:

The soil quality value is based on a random forest that was trained to predict expert knowledge based soil quality on a scale from 1 to 5 from the [SoilGrids250 data set](#) and the [WISE data set](#). The most influential variables in the random forest were total depth to bedrock and sand content. Thus the variable is a good predictor for plant available water storage capacity. This information has been provided in the protocol of this study.

3.3.31 COMMENT 35

L247 “The exact length of the simulation was also decided by the modeling teams” seems odd not to constrain that.

Response to comment 35:

We understand the reviewers concern on not constraining the length of the simulations, but we decided to leave it to the assessment of the individual modelers, as we wanted to obtain simulation that were run from bare ground in the absence of management to a simulated equilibrium (“potential natural vegetation”) with current climate. Each modeler know best how to obtain this in each model and different model uses its own logic to get to the equilibrium that can require different lengths.

3.3.32 COMMENT 36

L250 “The simulations were run in the absence of management to a simulated equilibrium” The absence of disturbance and running to equilibrium set quite specific conditions; suggest this is emphasised more in abstract / intro / discussion.

Response to comment 36:

We agree with the reviewer that the simulations set up regarding absence of management and to a simulated equilibrium are quite specific. We found challenging to incorporate this in the introduction as we see this a more of a methodological specification, and in the abstract due to space constraints as we though is more interesting to focus in the results and meaning of our findings than in methodological aspects. However we currently have a lengthy discussion on the equilibrium assumption and also about the importance of considering disturbances, which we think it would be a great further research direction to better understand regeneration from models of forest dynamics.

3.3.33 COMMENT 37

L252-253 “This entails the assumption that (1) the observations from the forest reserves reflect no traces of forest management,” That’s seldom the case in Europe where the legacies of management are evident centuries on

Response to comment 37:

The reviewer is correct and management legacies might be evident for long periods of time. The data were collected in forest reserves where no management has taken place for long periods of time which makes our assumption of an equilibrium between forest properties and environmental drivers reasonable. This is the best data we have and we have addressed its limitations, as this one, in the discussion because we are also aware that is not ideal.

3.3.34 COMMENT 38

L275 “Each of the models reported the regeneration number by sampling 200 times in a 10-year interval” regeneration number = number of “recruits” entering the 7-cm size class?

Response to comment 38:

Following the reviewer’s comment we have changed this sentence L[290-291]

3.3.35 COMMENT 39

L295 “(2) regeneration species diversity” species diversity of recruitment

L295 “(3) regeneration mortality” mortality of recruits

Response to comment 39:

We have changed the terminology.

3.3.36 COMMENT 40

L296 “ingrowth gradients along the regeneration niches.” ?? not clear to me. How is regeneration niche defined??

Response to comment 40:

We defined the regeneration niche as the passing of a diameter threshold of 7 and 10 cm and we focused both in the width of the regeneration niche (i.e., in environmental space) as well as the intensity of the regeneration process (i.e., the number of ingrowth trees per area and per unit of time) were of interest. This is currently defined in the material and methods section and in the protocol of this study.

3.3.37 COMMENT 41

Equations 1-5. not convinced these formulae need to be shown as the Shannon index is very well known. Consider converting the Shannon index to the effective number of species, by taking the negative of the SI and exponentiating it. This is gaining ground in ecology as it's more intuitive to interpret than SI (see paper by Jost).

L1046 Such converting to effective no of species using Jost formula

Response to comment 41:

We agree with the reviewer that the Shannon index is very well known index and its equation as well. However we think is still important that we define how we did calculate the Shannon index based on the available data we had as we had some internal discussion on how to do this and the meaning of this index in this context, we thought the reader will benefit from a more detailed explanation. In any case the exact details of how we did calculate this index based on the simulation outputs are also available in the software supplementary materials, in the script figures.R L42-78. We also thank the reviewer for the suggestion of using the effective number of species instead of the Shannon index, it is indeed a good alternative, however we have decided to leave it with the SI.

3.3.38 COMMENT 42

L308 "Mortality in tree regeneration was assessed based on the ratio of regeneration between the 7 and 10" of recent recruits. Now I understand the 7 and 10 cm thresholds mentioned earlier; please clarify earlier mention.

Response to comment 42:

Following the reviewer's comment on the lack of clarity of the selected thresholds explanation we have changed at the beginning of the section simulation protocol the first sentence where we explain that we considered tree regeneration as the passing of a breast height diameter threshold of 7 and 10 cm. We have also reported later on in this section (but before this line) that each of the models reported the number of trees crossing the two diameter thresholds L[290-291].

3.3.39 COMMENT 43

L311-313 "The Reineke self-thinning rule is usually calculated for even-aged, single species stands and assumes a fixed relationship between the number of stems and the quadratic mean diameter in fully stocked pure stands." more to the point, it's always (I think) applied to all stems in the stand, not just regenerating trees, as it's fundamental concept is that entire canopies are space filling . Why would we expect it to apply to a narrow cohort? I think not.

equation 5 Suggest writing N7 / N10

Given my comment above, is there a strong justification for the -1.605 exponent? Or just keep it at one?

Response to comment 43:

Following the reviewer recommendation we have changed the equation formulation to better represent the Reineke relationship. The reviewer is correct that the Reineke is applied to all stems in the stand, in our case the N value is only changing due to the available regeneration at the two diameter thresholds as the adult trees number is the same. By comparing our expected relationship of recruited trees at 7 and 10cm with the Reineke value we are just comparing to a theoretical value in fully stocked stands. We have used the coefficient -1.605 because Reineke attributed a general validity to this allometric coefficient for fully stocked, even-aged forest stands, regardless of tree species and site.

3.3.40 COMMENT 44

L317 “model type (empirical or process based),” if statistical analyses are used to estimate the coefficients of complex functions that underpin a process based model, is that model empirical or process based?

Response to comment 44:

The reviewer’s comment is a very valid one as it is true that many processes in process based models have been defined based on empirical approaches. In this study we have categorized the models in process based and empirical based on their general approaches to define forest dynamics. This categorization is available in table 1. It is also possible to find further details on each of the models in the supplementary materials.

3.3.41 COMMENT 45

L319-321 - overly complex sentence structure

Response to comment 45:

We have improved this sentence L[334-337].

3.3.42 COMMENT 46

L323-324 presumably, recruits make up a very small proportion of stand basal if the models have been run to equilibrium?

Response to comment 46:

Yes the reviewer is correct, the recruits represent a small proportion of the total stand basal area.

3.3.43 COMMENT 47

L327 of each species?

Response to comment 47:

Added as suggested L[343].

3.3.44 COMMENT 48

L332-335 This give a matrix of species-level variation in recruitment across sites, but how does this relate to the regeneration niche concept of Grubb 1977?

Response to comment 48:

We have changed the text to clarify this point. L[350-351].

3.3.45 COMMENT 49

L496 “confronted with a unique dataset from unmanaged”

L498 “The EuFoRIa data (Käber et al. 2023) are exceptional, particularly”

I think there are plenty of other datasets that could be used and have better information on smaller trees than EuFoRIa, so suggest a more nuanced sales pitch here.

Response to comment 49:

We agree with the reviewer’s comment that there are other forest inventory data networks outside Central Europe even on forest reserves that could provide this sort of data. However, within Europe no such data existed before, e.g. from unmanaged forests and collecting such a large number of records.

3.3.46 COMMENT 50

L525-527 “There are multiple constraints to the regeneration niche of tree species (Price et al. 2001), and therefore the absence of regeneration is likely to be common (Fortin and DeBlois 2007), even over larger areas” YES - good point. This should be featured more prominent I think

Response to comment 50:

Thank you for this remark, we agree with the reviewer that this is an important point and specially how different models are capturing or missing to represent the absence of regeneration. We have now captured this aspect also in the introduction in the sentence, L[147-148], on top of the current whole paragraph on this issue present in the discussion.

3.3.47 COMMENT 51

L530-533 “This substantial difference may be due to the fact that the simulation results were drawn from equilibrium forests, whereas in reality many of the forest reserves are recovering from past management activities and have become denser over the past decades (e.g., Heiri et al. 2009), leading to less regeneration than in an equilibrium situation.” But also --- browsing, establishment sites, competition with herb layer all reduce recruitment

Response to comment 51:

We have included this point by adding into this paragraph a new sentence: L[544-546].

3.3.48 COMMENT 52

L552 “there is no evidence that models with feedback from the canopy captured better the species” could you explain what you mean by “feedback” here?

Response to comment 52:

We have modified this sentence to clarify what do we mean here with feedback L[563-567]

This had also been defined in the material and methods where we stated L[180-182].

3.3.49 COMMENT 53

L586-587 “Our study showed that increasing complexity in the regeneration modules is not linked with a higher accuracy of the projections of regeneration levels,” Surely more complex models WOULD be better if they captured the relevant processes! Arguing for simple vs complex models is a distraction here... we need models that are as simple as possible but not too simple..

Response to comment 53:

We agree with the reviewers statement that we need the “right amount” of complexity which considers the most relevant processes. However there are limited studies looking at this aspect in models of forest dynamics and even more limited looking at this aspect in the regeneration related processes in models of forest dynamics. We still think it is important to discuss this issue in the context of our findings even though this was not the solely focus of our research.

3.3.50 COMMENT 54

L596-598 “Competition for light as a strong filter for tree regeneration has been widely documented (Messier et al. 1999; Collet and Chenost 2006; Berdanier and Clark 2016), but the models examined here did not reproduce this expectation.”, because ... remind us of the evidence you are putting forward

Response to comment 54:

Following the reviewers comment we have added the a sentence in this part of the text L[611-612].

3.3.51 COMMENT 55

L603-606 “This made it impossible to evaluate the regeneration for the extremes of the stand density ranges in some models. For example, regeneration levels at low stand densities are relevant to assess how well forests are recovering e.g. after gap creation due to disturbance” Agreed ... and this paper isn’t testing performance of models in predicting regeneration in large gaps. It’s possible that the models do a better job at this, but it’s not something you tested. Suggest making that point e.g. in the abstract.

Response to comment 55:

We have decided not to include this aspect in the abstract due to lack of space.

3.3.52 COMMENT 56

L656 “Yet, the global models should not be at a disadvantage due to the limited spatial consideration” Don’t quite follow your point here. what does “at a disadvantage” mean?

Response to comment 56:

In this sentence we are discussing how the sampling strategy presented in the study protocol might have represented a disadvantage for landscape level models due to the limited spatial scale, but how this might not be the same for global level models as they usually lack dispersal between cells.

3.3.53 COMMENT 57

L676-677 “Therefore, we further recommend that more effort should be invested into collecting harmonized datasets on tree regeneration.” regeneration in what sense? Seedlings/ ground layer/ What is really needed?

Response to comment 57:

We have changed this text L[687-688].

3.3.54 COMMENT 58

L685-687 “However, this will require an entirely different set of observed data, and potentially not all models of forest dynamics would be able to assess the relationship of these aspects on tree regeneration, e.g. due to the lack of disturbance or appropriate management modules.” more akin to data already available in national inventories I think

Response to comment 58:

We agree with the reviewer that a more systematic kind of inventory could be useful to detect regeneration patterns under management or disturbances impacts.

3.3.55 COMMENT 59

L696 “Exercises like the one presented here, where the models are operated in ”blind flight” mode.” Definition of blind flight “To do something based on guesswork, intuition, or without any help or instructions”....is that what you mean?

Response to comment 59:

With blind flight we refer to doing the simulations without having important information about the expected regeneration values.

3.3.56 COMMENT 60

Table 1 caption: A useful table. I'd encourage the authors to write a lengthier table heading, which briefly summarises the methods section, so that readers can know what "start from", "runtime for sampling" mean without having to delve into the text. Also I'd like to see an explanation of "formulation complexity" in the text and here. The supplementary information gives more details about model assumptions, and I do wonder whether those could be included here; in particular, seed production and dispersal is clearly important for regeneration, so could columns be introduced to compare assumptions of model with respect to these. Also deer browsing, herb-layers and nurse logs all have an influence of regeneration, but not incorporated into these models; I suggest the table heading makes that clear.

Response to comment 60:

Following the reviewers comment we have expanded the caption of table 1. We have decided not to include further variables on the model assumptions as we believe the current information provides a good general overview of the models approach and further details are included in the appendix.

3.3.57 COMMENT 61

L831 Add publisher

L939 Correct doi

Response to comment 61:

Changed as suggested.

3.3.58 COMMENT 63

L1041 suggest sticking with "ingrowth" (as used in the figure) or "recruitment"

Response to comment 63:

We have changed the terminology.

3.3.59 COMMENT 64

L1050 Overpredicted

Response to comment 64:

Changed as suggested.

For Review Only

Tree regeneration in models of forest dynamics: a key priority for further research

Olalla Díaz-Yáñez^{*1},

Yannek Käber¹,

Tim Anders²,

Friedrich Bohn³,

Kristin H. Braziunas⁴,

Josef Brůna⁵,

Rico Fischer³,

Samuel M. Fischer³,

Jessica Hetzer²,

Thomas Hickler²,

Christian Hochauer⁶,

Manfred J. Lexer⁶,

Heike Lischke⁷,

Mats Mahnken⁸,

Paola Mairota^{9₈},

Ján Merganič^{10₉},

Katarina Merganičová^{10,11,12},

Tobias Mette^{13₁₂},

Marco Mina^{14₁₃},

Xavier Morin^{15₁₄}

Mats Nieberg^{15_i}

Werner Rammer⁴,

Christopher P.O. Reyer^{8,15},

Simon Scheiter²,

Daniel Scherrer¹⁶,

Harald Bugmann¹

¹ETH Zurich. Department of Environmental Systems Science. Institute of Terrestrial Ecosystems. Universitätstrasse 16, 8092 Zürich, Switzerland. ²Senckenberg Biodiversity and Climate Research Centre (SBiK-F), 60325 Frankfurt am Main, Germany.

³Department of Computational Hydrosystem Ecological Modelling, Helmholtz-Centre for Environmental Research - UFZ, Permoserstraße 15, 04318 Leipzig, Germany. ⁴Ecosystem Dynamics and Forest Management Group, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany. ⁵Department of Geoecology, Institute of Botany of the Czech Academy of Sciences, Průhonice, Czech Republic. ⁶Institute of Silviculture, Department of Forest and Soil Sciences, University of Natural Resources and Life Sciences, Peter Jordan Strasse 82, A-1190 Vienna, Austria. ⁷Dynamic Macroecology, Swiss Federal Institute for

Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland. ⁸Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, 14473 Potsdam, Germany. ⁹Department of Soil Sciences, of Plants and Food,

University of Bari Aldo Moro, Italy. ¹⁰Department of Forest Harvesting, Logistics and Ameliorations, Faculty of Forestry, Technical University in Zvolen, Slovakia. ¹¹Czech

University of Life Sciences Prague, Faculty of Forestry and Wood Sciences, Kamýcká 129, 16500 Praha 6 – Suchdol, Czech Republic. ¹²Department of Biodiversity of Ecosystems and Landscape, Slovak Academy of Sciences, Štefánikova 3, P.O.Box 25, 81499 Bratislava, Slovakia. ¹³Dpt. of Soil and Climate, Bavarian State Institute of Forestry, Germany.

¹⁴Institute for Alpine Environment, Eurac Research, Bolzano/Bozen, Italy. ¹⁵CEFE, CNRS Campus du CNRS 1919, route de Mende 34293 Montpellier cedex, France. ¹⁶Swiss Federal

Institute for Forest, Snow and Landscape Research WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland. * Corresponding author (olalla.diaz@usys.ethz.ch)

OPEN RESEARCH STATEMENT: This submission uses novel code, both the data and code are provided in an external repository to be evaluated during the peer review process. The repository can be downloaded from [this figshare private link](#):

<https://figshare.com/s/100857ba67d2351dcb15>. The protocol is available at Díaz-Yáñez, Käber, and Bugmann (2022) (with link: <https://osf.io/czdxp/>). We will make publicly available the code and data if the manuscript is accepted.

Keywords: tree regeneration | tree recruitment | tree establishment | forest dynamics | models | landscape | stand | global

Abstract

Tree regeneration is a key process in forest dynamics, particularly in the context of forest resilience and climate change. Models are pivotal for assessing long-term forest dynamics, and they have been in use for more than 50 years. However, ~~an assessment of there is a need to evaluate~~ their ~~ability~~capacity to accurately represent tree regeneration ~~is lacking~~. We assess how well current models capture the overall abundance, species composition, and mortality of tree regeneration.

Using 15 models built to capture long-term forest dynamics at the stand, landscape, and global levels, we simulate tree regeneration at 200 sites representing large environmental gradients across Central Europe. The results are evaluated against ~~comprehensive~~extensive data from unmanaged forests.

Most of the models overestimate ~~regeneration~~recruitment levels, which is ~~only~~ compensated only in some models by high simulated mortality rates in the early stages of individual ~~tree~~tree dynamics. Simulated species diversity of ~~regeneration~~recruitment generally matches ~~the~~ observed ranges. Models simulating higher stand-level species diversity ~~at the stand level~~ do not feature higher ~~regeneration~~species diversity. in the recruitment layer. The effect of light availability on ~~regeneration~~recruitment levels is captured better than the ~~effect~~effects of temperature and soil moisture, but patterns are not consistent across models. Increasing complexity in the tree regeneration modules ~~of the models~~ is not related to higher accuracy of simulated tree ~~regeneration~~recruitment. Furthermore, individual model design is more important than scale (stand, landscape, global) and approach (empirical, process-based) for accurately capturing tree regeneration.

Despite ~~considerable~~the mismatches between simulation results and data, it is remarkable that most models capture the essential features of the highly complex process of tree regeneration, while not having been parameterized with such data. We conclude that much can be gained by evaluating and refining the modeling of tree regeneration processes. This has the potential to render long-term projections of forest dynamics under changing environmental conditions ~~that are~~ much more robust.

Introduction

Forests provide a wide range of products and services of vital importance to humankind (FOREST EUROPE 2020). Under the pressure of climate change, increasing disturbance impacts and changing societal demands on forest ecosystem services, it is becoming ever more important to understand how forest structure, composition and function will change, and to evaluate forest capacity to adapt to or be resilient after disturbance (Lindner et al. 2010; Seidl and Turner 2022). A wide range of models of forest dynamics ~~were have been~~ developed over the past decades considering the impacts of climate (Bugmann and Seidl 2022). From these studies, it is evident that we have a reasonably good understanding how to model tree growth (Vanclay and Skovsgaard 1997; Bugmann et al. 1996), and substantial efforts have been dedicated to improving the representation of tree mortality (Bugmann et al. 2019; Cailleret et al. 2017). In contrast, tree regeneration is ~~a~~ much less studied process ~~that and~~ is often represented rather coarsely in models (Walck et al. 2011; Leishman et al. 1992; Price et al. 2001). ~~This presents an important~~ research gap, particularly in the context of climate-induced forest disturbances and forest resilience.

Tree regeneration arises from multiple processes ~~such as including~~ pollination, fruit maturation, seed production, dispersal, germination, juvenile growth and survival (Price et al. 2001; Vacchiano et al. 2018). All these processes are difficult to assess, and some of them are scarcely understood and thus appear highly stochastic (Bogdziewicz et al. 2021). Identifying the appropriate level of complexity for the mathematical formulation of the key factors that are leading to successful tree regeneration is challenging. Currently, tree regeneration processes in dynamic forest models are handled in a multitude of ways (König et al. 2022; Bugmann and Seidl 2022): from 1) entirely ignoring ~~it them~~ (as done in classical forest growth models, e.g., Pretzsch et al. 2002), across 2) the use of a few simple environmental filters, as done in most forest “gap” models (Shugart 1984) and Dynamic Global Vegetation Models (e.g., Hickler et al. 2012; Smith, Prentice, and Sykes 2001a), to 3) complex approaches that incorporate local feedback from the canopy, multiple ecological processes and often also short time steps (e.g., Seidl et al. 2012; Wehrli et al. 2006), or 4) field-based statistical parameterizations, which however are not easy to extrapolate in space and time (e.g., Ribbens, Silander, and Pacala 1994).

Overall, models are needed to (1) synthesize existing empirical data and explore their relationships, (2) assess future tree regeneration e.g. in the context of global change scenarios, and (3) identify the most important processes that are shaping ~~ecological vegetation~~ patterns. Given the ~~current~~ strategies that are used in models of forest dynamics to represent tree regeneration, ~~model their~~ behavior often is prone to problems, such as very high levels of tree regeneration that necessitate excess mortality at early stages of tree life to simulate correct stand structure and composition. (Kroiss and Hille Ris Lambers 2015). Also, inadequately high species diversity in tree regeneration may be simulated, which is characteristic of “classical” forest gap models (Gutiérrez, Snell, and Bugmann 2016), at least as long as the simulation set-up comprises a multitude of species. Some models use calibration against local canopy-level data to constrain simulated ~~tree~~ regeneration ~~levels~~, which is likely to hamper the general applicability of these models e.g. under scenarios of climate change. Furthermore, correctly capturing the species

composition ~~inof~~ tree regeneration as a function of the presence of seed trees in the canopy is often a particular challenge, potentially leading to unrealistic successional drift in the model, which must be corrected by factors that are hard to parameterize (e.g., Lischke and Löffler 2006). A related issue is the excessive reduction of species diversity due to positive feedback effects, such that eventually just single-species stands remain. (Meier et al. 2011). This is sometimes corrected by the incorporation of a low level of seed influx of all species at all times (Schumacher et al. 2006) or by restricting the number of seeds per species in the seedbank (Lischke and Löffler 2006). However, simulated species composition is usually exceedingly sensitive to assumptions about seed availability, whereas the parameters of such functions are poorly constrained by field data. Lastly, there is often a problem with insufficient observational constraints on parameter values for models that start from very small tree sizes (e.g., 10 cm ~~ofin~~ height, or even ~~seedsfom seed~~) and track tree development in a process-oriented manner by considering a multitude of ecological influences, rather than emphasizing tree ~~regenerationrecruitment~~ into a larger size class (e.g., Wehrli et al. 2006).

~~Modeling tree regeneration processes is challenging, and even in empirical ecology, it has not received much~~ Thus, a focus on the modeling of tree regeneration processes is sorely needed and overdue (Price et al. 2001; Walck et al. 2011) if we are to make reliable projections of future forest dynamics, i.e., when the models need to be operated in extrapolation mode, as well as from a fundamental ecological point of view for increasing systems understanding. Also in empirical ecology, tree regeneration has not received a lot of attention (Hanbury-Brown, Ward, and Kueppers 2022), although many sub-processes have been studied in detail (Miina, Eerikäinen, and Hasenauer 2006). Yet, little data are available that covers all the processes within one species along environmental gradients, let alone for a vast suite of species. Thus, the available data are Data on forest regeneration are often fragmented, which constitutes a major problem for model building (Clark et al. 1999 Clark et al. 1999). On the one hand For example, monitoring on permanent plots (such as National Forest Inventories) often measures tree regeneration the ingrowth of new trees into a specific size class (Zell et al. 2019), however with a design that captures the rate after the stand initiation phase (Hallsby et al. 2015). Data from permanent plots, although available across regions, are also highly heterogeneous, rendering their use difficult in a modeling context (Käber et al. 2021). On the other hand Lastly, targeted studies to measure tree regeneration on experimental sites are often limited in spatial extent, species studied, or the sub-set of processes that are targeted investigated (Berdanier and Clark 2016; Collet and Chenost 2006).

~~Thus, a focus on the modeling of tree regeneration processes is sorely needed and overdue (Price et al. 2001; Walck et al. 2011) if we are to make reliable projections of future forest dynamics, i.e., when the models need to be operated in extrapolation mode, as well as from a fundamental ecological point of view for increasing systems understanding.~~ In the present study, for the first time, a large number of forest models commonly used to assess forest dynamics under climate change are evaluated against a continental-scale, multi-species harmonized dataset on tree ~~regenerationrecruitment~~ (Käber et al. 2023). By tree ~~regenerationrecruitment~~, we refer to the passing of trees across a specific diameter threshold (“ingrowth”). We included models that are based on a range of “philosophies”,

from models operating at the stand to the global scale as well as the range of models from empirically derived to “process-based” (e.g., Bugmann et al. 1996; Fabrika M. and Ďurský 2012; Lexer and Hönninger 2001; Reyer et al. 2014). Due to the large variability in tree regeneration patterns in nature and the large number of factors driving and constraining this process — including some that are not incorporated explicitly in most models, such as deer browsing — we do not aim for a detailed statistical evaluation of each model. Instead, we aim to evaluate the general regenerationrecruitment patterns and magnitudes simulated by the models, and benchmark the simulated regeneration niche of multiple species against empirical data along a wide environmental gradient of temperature, moisture, and light availability (Grubb 1977).

More specifically, we aim to answer the following questions: (1) Are models of forest dynamics capturing accurately tree regenerationrecruitment levels, initial tree species diversity and mortality in the regenerationrecruitment? (2) Do model traits explain differences in model performance? (3) How well do the models capture total regenerationrecruitment and individual speciesthe regeneration niches of individual species across environmental gradients of light availability, temperature, and soil moisture?

We evaluate the performance of the models by analyzing simulated data alongside observed data. Specifically, we examine tree species diversity, mortality rates in tree recruitment, and recruitment levels. Then, we link model traits (e.g., complexity, scale) to performance, and we assess whether modeled total recruitment or individual species regeneration niches (i.e., proportional recruitment along an environmental gradient) align with observational data.

Material and methods

Models

Fifteen models of forest dynamics (including two models featuring model variants) developed for the stand, landscape or global scale were used to simulate forest dynamics ~~at the stand, landscape or global scale~~ (Table @ref(tab:tableModels)). The approaches used for model construction and their origin differ strongly, with most of the models featuring a largely “process-based” approach, whereas two models are based on formulations derived from the statistical analysis of inventory data (SIBYLA and xComp). Some of the models largely rely on the original approach underlying forest gap models (e.g., ForClim 1), whereas some are based on plant ecophysiological processes (e.g., FORMIND and iLand). The different approaches underlying the models have strongly influenced the formulation of tree regeneration processes.

The regeneration models can be differentiated into ‘regeneration’ and ‘recruitment’ approaches (Vanclay and K. 1994; König et al. 2022). *Regeneration models* include processes such as flowering and pollination, seed production, seed dispersal, germination, and seedling growth, which ultimately lead to the simulated number of established trees. *Recruitment models*, in contrast, introduce a number of new trees with certain

characteristics such as biomass or diameter, without explicitly considering earlier development processes. We can further distinguish models that feature a feedback in the regeneration from stand properties to simulated recruitment, i.e., where the level and species composition of regeneration recruitment is influenced by the existing adult tree community via the production of seeds, (or seedlings or saplings;), from models that do not contain such-a feedback.

The starting point for tree regeneration in the models differs as well, ranging from seedbank, seed or seedling to sapling (i.e., trees much larger than 10 cm height, often ca. 2-3 m tall). Models that start from seed need to include a larger number of ecological processes such as germination and survival of young seedlings, whereas models that start from saplings have to aggregate via parameterizations several ecological processes that are not treated explicitly. This latter approach reduces model complexity but comes at the cost of blurring blurred process representation.

In the models used here, the overall complexity in the regeneration modules varies considerably. Following Bugmann and Seidl (2022), we can classify seven models (iLand, PICUS, LandClim, ForCEEPS, LPJ-GUESS, ForClim and TreeMig) as having rather high complexity in their regeneration modules (mean regeneration formulation complexity across all processes >0, Table @ref(tab:tableModels)).), whereas the other eight models fature relatively low complexity. Two models, ForClim and ForCEEPS, were used here with two alternative variants of regeneration. ForClim variant 1 (Bugmann et al. 1996) is based on a recruitmentregeneration module that adheres closely to the concept introduced by Botkin, Janak, and Wallis (1972), whereas) which is based on environmental filters considering species-specific thresholds of light availability and climatic variables. In contrast, ForClim variant 11 is adopting a different slightly more complex approach where individual species properties and their relation to the environment are importantrepresented gradually along with the relative suitability of different species (i.e., competition among tree species) (Huber, Bugmann, and Lafond 2020). These two model variants allow us to evaluate a more process-based and complex module (variant 11) against a simple module (variant 1), while the rest of the model structure is identical. Similarly, the two ForCEEPS variants allow us to isolate the importance of the canopy feedback (i.e., simulated actual composition and relative abundance of species in the plot) via seed trees for the quantity and quality (e.g., diversity and composition) of simulated regeneration, as one variant includes this feedback whereas the other does not.

Observed data

RegenerationRecruitment data covering a wide range of environmental conditions are hard to obtain, and this is one of the reasons why most models of forest dynamics have never been confronted with a dataset covering such gradients over a large number of sites, to evaluate how well regeneration is captured. The observations used here are derived from a novel and unprecedented-network of sites in forest reserves that represent the range of environmental gradients in temperature and precipitation in Central Europe as compiled in the framework of the EuFoRia network (EuFoRia 2019) (Figure @ref(fig:map); Käber et al. 2023) (Figure 1). The environmental coverage and the variety of forests represented by this data is unprecedented within Europe. These forests have been unmanaged for at least

10 years prior to their designation as reserves, and most of them provide time series of natural forest dynamics over multiple decades and up to ca. 8085 years. The census periods range from 3 to 37 years, with an average of 14 years. The data provide information at the tree level, thus allowing for the sequential comparison of processes such as individual-tree regeneration recruitment and death between the measurements.

We selected 200 sites from this network as the benchmarking dataset for the simulation, so as to be representative of the environmental variation contained in the data. This was achieved by applying k-means clustering to define 200 clusters of plots from the original set of 869 plots along the environmental dimensions of temperature, climatic water balance, soil quality and exposition (i.e., slope and aspect). Each of these sites featured at least two consecutive measurements. Regeneration Recruitment size thresholds for these sites differed between ranged from stem diameters of at breast height (dbh) from 0 and to 10 cm. For our study, we defined two datasets, where one included 165 sites with a diameter threshold of 7 cm or lower, and the other included another 35 sites with diameter thresholds between 7 and 10 cm. Therefore we had in total 200 sites with observations of newly recruited trees above 10 cm and a subset of 165 sites with recruitment data above 7 cm. In the results section, any observations from the data at the 7 cm threshold originated from 165 sites.

Plot size ranged from 0.02 ha to 5.52 ha, and the data were further processed and aggregated following Käber et al. (2023) to provide regeneration recruitment rates per ha and per decade. The observations used in this study featured 30,900 newly established trees. Regeneration Recruitment rates per site, sample and decade ranged from zero to 12461,246 trees, with a mean of 56 trees. Adult species composition was also available for each plot. For more details on this unique dataset and the detailed regeneration recruitment information, cf. Käber et al. (2023).

Simulation protocol

The overarching goal of the simulation experiments was to assess the tree regeneration recruitment as it arises from empirical data against its representation in a wide range of models of forest dynamics. We define tree regeneration recruitment as the passing of a breast height diameter threshold of 7 or and 10 cm, respectively (synonym: ingrowth). To this end, each modeling group was provided with a detailed protocol (Díaz-Yáñez, Käber, and Bugmann 2022) with instructions how to perform the simulations, providing input variables on climate and soil conditions, and the list of expected output variables. Neither were further site information (except for the data specified below) nor any data on tree regeneration recruitment or forest stand features provided prior to the simulation. That is, the models were run in “blind flight” mode.

The input variables were collected from different data sources and aggregated to be adapted to suitable for the needs of the different models. Time series of climatic variables from 1981 to 2018 were provided in hourly (Era5-land data, Muñoz Sabater 2019), daily and monthly resolution (CHELSA data, Karger et al. 2021). Some variables required for some models, such as relative humidity or vapor pressure deficit, were calculated from these variables. The final instructions for using the climate data were slightly different in

each model, based on the approach that best suited the model (Table @ref(tab:tableModels)). Soil quality data were provided as continuous values between 1 and 5 (Soilgrids dataset, Hengl et al. 2017). The protocol also provided the elevation, slope and aspect for each of the 200 sites (ASTER Science Team 2019), but no other spatial information such as coordinates, with the exception of iLand and aDGVM2, which required blurred coordinates to derive highly detailed soil data.

The simulations were run in the absence of natural disturbances. Only the model LPJ-GUESS had to include a “background” disturbance probability to increase chances of shade-intolerant species to establish. The simulations were set up to sample species-specific regenerationrecruitment rates per decade and per ha in the equilibrium state of the model, typically entailing a “spin-up” run (as we did not provide any forest data). The modeling teams decided on the simulated area and how they derived these samples. The exact length of the simulation was also decided by the modeling teams (Table @ref(tab:tableModels)). Further details on how each modeling team prepared the simulations and the outputs are available in Supplementary Material 1.

The simulations were run in the absence of management to a simulated equilibrium (“Potential Natural Vegetation”) with the current climate, as the ultimate goal was to evaluate tree regenerationrecruitment under comparable and near-equilibrium conditions. This entails the assumption that (1) the observations from the forest reserves reflect no traces of forest management, and (2) there is an equilibrium between forest dynamics and climate. While the former might be starting to be visible in many of the EuFoRIA reserves, the latter may be more debatable. However, in the absence of detailed data on the history of each plot in the EuFoRIA network, some broad assumptions had to be made. Both the width of the regeneration niche (i.e., in environmental space) as well as the intensity of the regeneration processrecruitment (i.e., the number of ingrowthrecruited trees per area and per unit of time) were of interest.

The simulations were run for mixed-species forests (not multiple single-species simulations) using mixtures of eleven species or genera for which regeneration data of sufficient quality were available from EuFoRIA: *Fagus sylvatica* L., *Picea abies* L., *Abies alba* Mill., *Carpinus betulus* L., *Tilia cordata* Mill., *Acer pseudoplatanus* L., *Betula* spp. L., *Fraxinus excelsior* L., *Quercus* spp. L., *Alnus glutinosa* L., and *Pinus sylvestris* L. The same set of species was used at all 200 sites. Two models included their standard set of species for the simulations, which is much larger (i.e., ForClim 1, ForClim 11, TreeMig~~1~~), and the species simulated beyond the eleven species were lumped in a category “others”. In three models, fewer than these eleven species were simulated (4C, xComp, LPJ-GUESS) (Table @ref(tab:tableModels)). In 4C only *Fagus sylvatica*, *Picea abies*, *Betula* spp., *Quercus* spp. and *Pinus sylvestris* are parameterised. xComp simulations did not consider *Fraxinus excelsior* due to a deprecated species parameterisation. In LPJ-GUESS, *Acer pseudoplatanus* and *Alnus glutinosa* are not parameterized and therefore these species could not be included in the simulation. Finally, the model aDGVM2 does not simulate individual species. Rather, community assembly processes and trait filtering generate plant communities that are adjusted to the local biotic and abiotic conditions, and the simulated plants can be classified into ecological strategies based on their trait values in a post-processing step.

Each of the models reported the regeneration number of new trees crossing the two diameter thresholds by sampling 200 times in a 10-year interval for each species and per ha for each of the 200 sites. Multiple samples per site were used to better understand the simulated variation within each site. This was done using different strategies, depending on the model, including (1) sampling simulated data from the same 1 ha plot in the equilibrium over time, (2) sampling several 1 ha plots from the simulated forest at one specific point in time (in the equilibrium) or (3) a combination of (1) and (2). This resulted in 880,000 observations per model that simulated the 11 species included in the protocol (200 sites, 200 samples per site, 11 species and two diameter thresholds). For the models that simulated additional species, their regeneration rates were aggregated as “others”.

Two models did not provide results from all the simulated samples or sites to avoid unrealistic results, as follows. In the model 4C, a threshold of a maximum basal area of 90 $m^2 ha^{-1}$ was used to avoid unrealistic stand basal area data, and therefore not all the samples and sites were reported. The reason for this is that 4C is not suitable for long-term simulations without management, due to misrepresentations in density-dependent mortality processes in long-term simulations and assumptions of tree geometry that lead to unrealistic single tree dimensions of very old individuals. LPJ-GUESS had 2% of the sample outputs without tree regeneration recruitment, and these were considered as zero stand basal area and zero regeneration for all the tree species simulated; one site produced grassland rather than a forest and was not included in the results.

Data analysis

The simulation results were analyzed regarding (1) regeneration recruitment levels (i.e., ingrowth number per unit time and space), (2) regeneration recruitment species diversity, (3) regeneration recruitment mortality, (4) the relationship between model performance and model traits, and (5) ingrowth gradients along the regeneration niches. We evaluated species diversity across the models and in relation to the observed data by calculating the Shannon index H based on the relative proportion of the species in terms of basal area. It was calculated for the regeneration (H_{R_n}) (Equation) and at the stand level ($H_{S_{n,s}}$) (Equation). The higher the value of the index, the higher the species diversity at a particular site and sample.

where s is the total number of species present and that have a basal area larger than zero in sample n ; $p_{R_{n,i}}$ is the proportion of species i in sample n calculated as the regeneration basal area ($r_{BA_{n,i}}$) for that species i relative to total recruited basal area ($TotalR_{BA_n}$) of the sample n ; $p_{S_{n,i}}$ is the proportion of species i calculated as the basal area of all trees ($BA_{n,i}$) of that species i relative to total basal area ($TotalR_{BA_n} Total_{BA_n}$) of the sample n . Species diversity was not assessed for aDGVM2 as this model does not simulate individual species.

Mortality in tree regeneration recruitment was assessed based on the ratio of regeneration recruitment between the 7 and 10 cm diameter thresholds. We used the Reineke self-thinning rule (Reineke 1933; Hans Pretzsch and Biber 2003) as a reference to estimate whether the ratio of regeneration recruitment between the 7 and 10 cm diameter these two thresholds were was above or below the expected theoretical rate. The

Reineke self-thinning rule is usually calculated for even-aged, single species stands and **assumesis based on** a fixed relationship between the number of stems and the quadratic mean diameter in fully stocked pure stands. The value used in our comparisons was 1.77 (i.e., we expect stem numbers at 7 cm to be 77% higher than at 10 cm), calculated using Equation .

We assessed model performance in relation to model traits focusing on (1) model complexity as defined by Bugmann and Seidl (2022), (2) model type (empirical or process-based), (3) the presence or absence of a canopy feedback for regeneration, and (4) the scale of application of the model (stand, landscape or global). We tested for significant differences **using two-sided t-tests by considering each of these model features regarding the observed and simulated regeneration levels and species diversity, respectively between simulated and observed values of the recruitment levels and species diversity by using t-tests with a correction for multiple testing. Note that the t-tests does not present conclusive evidence for a specific hypotheses, but rather facilitates managing the extensive number of comparisons discussed in our study.**

Total **regenerationrecruitment** and the regeneration niches of the individual species were evaluated across the **environmental** gradients of light availability, temperature, and soil moisture. **as captured in the data**. Stand basal area (the basal area of all the trees in each sample **including tree regenerationabove the recruitment threshold**) was used as a proxy for light availability at the forest floor, the annual degree-day sum (Allen 1976; Fischlin, Bugmann, and Gyalistras 1995) as a proxy for growing season warmth, and the climatic water balance as a proxy for soil moisture (Speich 2019). **RegenerationRecruitment** values **of each species** were calculated as the mean across the 200 samples per site. The observed data were modeled using a Generalized Additive Model (Wood 2011) with a negative binomial distribution and restricted maximum likelihood to better understand the relationship between the environmental gradients and the levels of observed regeneration, relative to the simulation results per model.

In order to analyze the regeneration niches across the climatic gradients, we focused on five common tree species or genera: *Abies alba*, *Fagus sylvatica*, *Picea abies*, *Pinus sylvestris*, and *Quercus* spp. For these species, we calculated the share in the **regenerationrecruited** basal area per site, **where each site represents a part of the environmental space**, as the mean across the available samples per site:

where s is the total number of species simulated, and $\bar{r}_{t,i}$ is the mean basal area in the **regenerationrecruitment** (subsequently referred to as “**regenerationrecruited** basal area”) of species i per t site across the available samples at that site. Furthermore, regarding the share in the **regenerationrecruited** basal area per species ($\bar{R}_{BA_{share_i}}$), we categorized this as zero when the **regenerationrecruited** basal area of that species ($\bar{r}_{BA_{t,i}}$) was zero, but also when both the **regenerationtotal recruited basal area** ($\sum_{i=1}^s \bar{r}_{BA_{t,i}}$) and the **recruited** basal area of that species ($\bar{r}_{BA_{t,i}}$) and the **total regeneration basal area** ($\sum_{i=1}^s \bar{r}_{BA_{t,i}} \bar{r}_{BA_{t,i}}$) equaled zero.

Results

RegenerationRecruitment levels

Simulated regeneration levels of recruitment varied strongly across the 15 models and typically did not match the levels found in the forest reserve data (Figure @ref{fig:over}). Regeneration2. Recruitment was overestimated in most models for both the 7 and 10 cm diameter thresholds, with the exception of the empirical stand model SIBYLA, the landscape model Landis-II and the global model aDGVM2, which estimated regenerationrecruited levels at the lower end of the plausibility interval of the observed data (Figure @ref{fig:over})-2. The models with the largest overestimation were the stand model PICUS and the landscape model TreeMig. For most models, the variability of simulated regenerationrecruitment levels across the 200 sites (visible from the interquartile range in the box plots of Figure @ref{fig:over})-2) was similar to or smaller than observed, with the notable exception of PICUS, where simulated regenerationrecruitment variability was much larger.

Both the observations and the simulated data had no regenerationrecruitment in some samples and at some sites (for details, cf. Table S1). The observed data had 4% of the samples with no regenerationrecruitment. Only three models had a larger proportion of no regenerationrecruitment (4C, Landis-II and aDVM2). Two models always simulated regenerationrecruitment for both the 7 and 10 cm threshold (xComp and TreeMig), i.e., they did not feature any zero values. The other ten models had a very low percentage of samples with no regenerationrecruitment (0.01% - 2.39%), i.e., they had distinctly fewer occurrences of zeros compared to the observations.

Tree species diversity of regenerationrecruitment

Most models matched the level of diversity of the observed data quite well (Figure @ref{fig:H710})-3. Five models overestimated regenerationrecruitment diversity: ForCEEPS, ForClim 1, PICUS, TreeMig and LandClim, the latter particularly for the 7 cm diameter threshold. The model 4C is a special case, as it simulated five species only, i.e., its diversity values are not directly comparable to those of the other models, nor to the observations. Only one model, Landis-II, consistently underestimated regenerationrecruited diversity.

In most models, there were only small but significant differences in the species diversity of regenerationrecruitment between 7 and 10 cm across sites (Tables S2). Four models (iLand, Landis-II, TreeMig and LPJ-GUESS) maintained the regenerationrecruitment diversity between the 7 and 10 cm thresholds (i.e., the differences between them were not significant, see Table S2), and the same was evident from the observed data.

For both the observed and the simulated data, and both regenerationrecruitment thresholds, species diversity in the regenerationrecruitment compared to stand-level diversity did not feature clear patterns across the models (Figures S1 and S2). Most models captured reasonably well or overpredicted species diversity at the stand level (Figure @ref{fig:combinedRichness})A4A, left and center; cf. Figure S3), but a characteristic feature

was that the simulations had a much lower variability of diversity compared to observations. Models overpredicting stand-level species diversity included ForClim1, ForClim 11, FORMIND and TreeMig (Figure @ref(fig:combinedRichness)A4A, left). Several models underpredicted species diversity, i.e., LPJ-GUESS, ForCEEPS(f) and 4C (Figure @ref(fig:combinedRichness)A4A, right).

The majority of the models overpredicted regenerationrecruitment diversity; as already observed in Figure @ref(fig:H710),3, only two of them underestimated it (4C and Landis-II) (Figure @ref(fig:combinedRichness)B4B and Figure S4). Both the models with feedback and those without feedback from the adult trees to regeneration (via seed production) overestimated, underestimated or captured reasonably well regenerationrecruitment species diversity compared to the observations. Again, simulated diversity had considerably lower variability than observed diversity in the regenerationrecruitment.

Mortality in tree regeneration

Recruitment mortality

There were strong differences among the models regarding the mortality rate between 7 and 10 cm (Figure @ref(fig:mortality11)).5. Most of the models featured a mortality rate significantly larger than the observed data, thus at least partially compensating for the general overestimation of regenerationrecruitment levels at 7 cm (Figure @ref(fig:over))2 and Table S3); six models did not feature a significant difference compared to the observed data (ForCEEPS, FORMIND, ForClim 1 and 11, iLand and aDGVM2). Four models (xComp, PICUS, LandClim and TreeMig) featured very high mortality rates (i.e., well above the Reineke self-thinning line), which compensated for the strong initial overestimation of regenerationrecruitment (cf.- Figure @ref(fig:over)).2. Interestingly, two models that underestimated overall regenerationrecruitment levels (SIBYLA and Landis-II; cf. Figure @ref(fig:over))2 featured mortality rates that were close to but still above the self-thinning line (Figure @ref(fig:mortality11)).5.

The fact that some models (and observed data) featured lower mortality than expected by Reineke's self-thinning rule may point to facilitation, or simply a higher mortality rate before the trees had reached 7 cm diameter. Yet, the case of models such as xComp, PICUS and TreeMig, whose mortality was well above the self-thinning line, indicates that they feature higher regenerationrecruitment mortality in mixed-species forests compared to even-aged, single-species forests. Some models yielded a mortality rate of regenerationrecruitment that is broadly compatible with the self-thinning rule (e.g., 4C and LandClim).

The three models with the largest overestimation of the proportion of regenerationrecruitment at the 7 cm threshold featured the highest mortality between the 7 and 10 cm thresholds (Figure @ref(fig:mortality13)A).6A). The other models that overestimated regenerationrecruitment had mortalities lower than the expected self-thinning ratio, with the exception of LandClim (Figure @ref(fig:mortality13)B).6B). It is noteworthy that most of the models that featured a low ratio of regenerationrecruitment between 7 and 10 cm (i.e., ForCEEPS, ForClim 1, ForClim 11 and iLand) had only a small

overestimation of regenerationrecruitment at the 7 cm threshold (Figure @ref(fig:mortality11)).5). aDGVM2 was the only model that underestimated regenerationrecruitment at the 7 cm threshold while featuring a low mortality between the two thresholds (Figure @ref(fig:mortality13)C6C), and only a few models were close to the correct regenerationrecruitment levels at 7 cm while also being close to the Reineke line (SIBYLA, Landis-II; cf. Figure @ref(fig:mortality13)D6D; and 4C as well as LandClim).

Lastly, some models that had a low number of individualsrecruits at the 7 cm threshold also had a smallersmall decrease of regenerationrecruitment between 7 and 10 cm (ratio 7-10 cm) (Figure S5). In these models, competition and self-thinning are either not pronounced or must have occurred before the trees had reached 7 cm. However, this pattern was not consistent across models (Figure S5). It is noteworthy that there is a relationship between the regenerationrecruitment levels at 7 cm and the mortality rate (Table S4): most of the models showed a positive effect (higher regeneration at 7 cm is coupled to an increase of the mortality between 7 and 10 cm). While some models showed a negative effect, the linear trend was not pronounced and not always significant regarding theits slope.

Model performance and model traits

There was no significant relationship between the mean complexity of the regeneration module (Table @ref(tab:tableModels)) and the overestimation proportion at a diameter threshold of 7 cm (Figure S6). All the models except 4C had significant differences between the observed and simulated mean regenerationrecruitment values. Therefore it was not possible to differentiate in terms of performance between models type (empirical or process based), or their scales (stand, landscape or global) (Table S5). The values of regeneration-species diversity in the recruitment simulated by each model were also assessed against the observed data to see if there was a difference between models that include a feedback compared to those that did not; all the models except ForCEEPs(f) and FORMIND had significant differences (Table S6). It is noteworthy that for the model ForCEEPs the pattern changed from the overprediction of species diversity in the regeneration to a diversity level that is closer to observations when the recruitment module included feedback (ForCEEPs(f)).

Regeneration gradients onand regeneration niches

When evaluating regenerationtotal tree recruitment levels along key gradients of light availability (basal area), temperature (degree-days) and soil moisture (climatic water balance), distinct features emerged: the models reproduced the effect of basal area in both its magnitude and patterns (Figure @ref(fig:nvTrend710BALightTemp)7 and S7) considerably better than the effects of the climatic gradients, where they featured varying patterns.

In the observed data, regenerationrecruitment levels decreased clearly with increasing stand basal area (Figure @ref(fig:nvTrend710BALightTemp)a7a). This trend was captured by the Landis-II model only, albeit at much lower values of basal area than in the observations. The other models featured distinctly different trends, such as (1) an increase

in regenerationrecruitment levels with increasing total basal area followed by the absence of regenerationrecruitment at high values of basal area (ForCEEPS, ForCEEPS(f), FORMIND, TreeMig and LPJ-GUESS), (2) almost constant regenerationrecruitment levels with basal area (ForClim 11, xComp, LandClim and aDGVM2), or (3) an increase of regenerationrecruitment up to a certain value of stand basal area followed by a decrease at even higher values of stand basal area, with model-specific thresholds (4C, ForClim 1, PICUS, SIBYLA and iLand).

The observed regenerationrecruitment did not change much across the climatic water balance gradient, and showed a slight increase with temperature. The Neither did the models did not generally match the observed regenerationrecruitment patterns across the these gradients of soil moisture and temperature (Figure @ref{fig:nvTrend710BALightTemp}b7b,c and S7), nor was there a clear pattern across models. Regarding the soil moisture gradient, a group of models featured decreasing regenerationrecruitment with increasing soil moisture (4C, ForClim 1, ForClim11, SIBYLA, xComp, PICUS, LandClim, TreeMig and iLand), while a few models showed the opposite trend (ForCEEPS(f), FORMIND and LPJ-GUESS). Across the temperature gradient, all the models that featured a decrease in regenerationrecruitment levels with increasing soil moisture showed the opposite trend (i.e., more regenerationrecruitment with increasing temperature), except for 4C. The other models did not feature a clear pattern.

Silver fir (*Abies alba*), beech (*Fagus sylvatica*), Scots pine (*Pinus sylvestris*) and oak (*Quercus* spp.) were the species observed most frequently in the EuFoRla data. Most models captured well the share of basal area in the regenerationrecruitment of these five species compared to the observed data along the two major climatic gradients (Figure @ref{fig:diff}, cf. Figure S8). However, the models differed considerably in the way in which they simulated these climatic niches.

First, some models overestimated regenerationrecruitment across the entire gradients of at least one of the five main tree species (Figure @ref{fig:diff}, e.g., ForCEEPS, ForCEEPS(f), LandClim or LPJ-GUESS), while other models overestimated regenerationrecruitment of more than one species for a large part of the climate gradients, such as xComp, PICUS, SIBYLA or LPJ-GUESS (cf. red colors in Figure @ref{fig:diff}). These trends were consistent for both regenerationrecruitment thresholds, although the exact changes in the share of the regenerationrecruited basal area were different (Figure S8).

Second, there was some congruence in the simulation results by species across the models. The share in the regenerationrecruitment of Silver fir (*Abies alba*) was captured well across the gradient by most models except for xComp, PICUS, and LandClim, which overestimated the regenerationrecruitment share of this species. There were also some peculiarities evident for some models, such as 4C that did not simulate *Abies alba*. The patterns across models for *Fagus sylvatica* regenerationrecruitment were more complex, as its regenerationrecruitment was underestimated by many models across the environmental gradients featuring negligible regeneration at most sites, while others overestimated it in different parts of the environmental space. Most models overestimated the share of *Picea*

abies regenerationrecruitment in the cold-wet part of the gradients. *Pinus sylvestris* and *Quercus* spp. represented a small share of the *regenerationrecruitment* in the observed data mostly at warm-dry sites. Many models, however, erroneously featured *regenerationrecruitment* for these two species along most of the climatic gradients, although with a low share of basal area. Five models (4C, FORMIND, SIBYLA, xComp, and LandClim) had *almost no regenerationvery small amounts of recruitment* of *Pinus sylvestris* and *Quercus* spp., or did not feature any *regenerationrecruitment* of these species at all.

Lastly, there is no model that performed well across the five species regarding the species-specific *regenerationrecruitment* levels in the environmental space (Figure @ref(fig:diff)).8). Some models (e.g., ForClim11, PICUS) tended to consistently overestimate the *regenerationrecruitment* share of all five species, but most models overestimated the *regenerationrecruitment* share of some species while underestimating the share of others (e.g., ForCEEPS, LPJ-GUESS).

Discussion

Tree regeneration is a fundamental process in forest dynamics. Correctly capturing *tree regenerationit* in dynamic models is fundamental to, for example, evaluate post-disturbance dynamics and potential long-term recovery trajectories as it will define the forest state in the first decades (Seidl and Turner 2022). If this initial phase is not captured well, we can not properly assess aspects such as forest resilience or the timing, magnitude and progression of carbon sequestration. In this study, for the first time, projections of tree *regenerationrecruitment* from multiple models of forest dynamics were confronted with a unique dataset from unmanaged forest reserves across a large environmental gradient in Europe.

The EuFoR1a data (Käber et al. 2023) are exceptional, particularly with respect to the number of records (number of sites and repeated measurements *on unmanaged forests*), which is essential for capturing a highly ‘noisy’ process such as tree regeneration. The use of this dataset for model benchmarking provided novel insights on the ability of state-of-the-art models to accurately simulate *regenerationrecruitment* levels, *its* species composition, and mortality in an early stage of tree life, i.e., between *a-tree diameterdiameters* of 7 and 10 cm. Overall, by adopting this approach a much broader understanding resulted than if we had used these data for model calibration: it is primarily from the shortcomings of the models that we can gain ecological insights (cf. Trugman 2022).

RegenerationRecruitment levels

Most of the models overestimated tree *regenerationrecruitment* levels. This has potentially far-reaching implications, e.g., regarding biomass (and thus carbon) turnover, with a potential overestimation of the capacity for forest carbon sequestration (Pugh et al. 2019; Pan et al. 2011). Yet, we focused on one specific stage, i.e., *regenerationrecruitment* into the 7 and 10 cm diameter classes. Trees at this size contribute little to carbon sequestration, and if the excess regeneration at this stage is compensated – in the models – soon

thereafter by higher mortality, overall simulated regeneration tree population dynamics may still yield trustworthy results at the level of stand structure and related forest attributes such as basal area or biomass.

From a modeling point of view, excess regeneration tree recruitment inevitably requires excess mortality rates in a later stage, either – as observed for some models – between the 7 and 10 cm diameter thresholds, or soon after the 10 cm threshold has been crossed. In any case, correcting at early stages for the expected forest densities at later stages is equivalent to compensating for a first error (excess regeneration recruitment) by a second error (excess mortality). It is highly likely that biased projections will result, because the two errors are unlikely to be perfectly linked and thus will not always compensate each other. Hence, this structural problem of most models investigated here is problematic particularly if the models are to be used under novel conditions such as under climate change (e.g., Huber et al. 2021) or in a decision support context (e.g., Thriplleton et al. 2021).

Regeneration Recruitment levels define the structure and composition of future forests, but it is equally important to correctly identify areas where regeneration is lacking (Rammer et al. 2021). There are multiple constraints to the regeneration niche of tree species (Price et al. 2001), and therefore the absence of regeneration is likely to be common (Fortin and DeBlois 2007), even over larger areas such as the one-hectare plots used here. Tree regeneration data are characterized by zero-inflation. This was clearly evident from the EuFoRIa dataset, but some of the models did not produce zeros at all, or featured a very low proportion of zero data. This substantial difference may be due to the fact that the simulation results were drawn from equilibrium forests, whereas in reality many of the forest reserves are recovering from past management activities and have become denser over the past decades (e.g., Heiri et al. 2009), leading to less regeneration than in an equilibrium situation. Another possible reason for this difference may be the exclusion of factors like deer browsing or failure to accurately measure establishment filters, such as competition with the herb layer, site-specific or environmental limitations.

Species composition of regeneration recruitment

Correctly capturing species composition in tree regeneration recruitment is important to assess the future functional diversity of a forest, e.g., its sensitivity to drought or resilience to disturbances (Redmond et al. 2015; Seidl and Turner 2022). In the simulations, overall species diversity levels in the regeneration recruitment were well within the observed range for half of the models (7 out of 14). Thus, while most of the models are facing difficulties to quantitatively match regeneration recruitment levels (cf. above), their performance is better when it comes to matching regarding the composition of regeneration recruitment as a function of abiotic and biotic conditions. Most models maintained or even decreased regeneration recruitment diversity between the 7 and 10 cm thresholds, and the same is visible from the empirical data, but the differences between the 7 and 10 cm thresholds were not significant. Diversity variations may be explained by the way the individual models consider regeneration processes (König et al. 2022), e.g., without considering the feedback from adults, regeneration can be linked to either the

seed/seedling bank, seed rain from dispersal or from a seed/seedling bank; neither process is incorporated across many models studied here, however pool of available species.

The species diversity of the entire stand was captured better by the models than the diversity of tree regeneration recruitment, and only a few models overpredicted stand-level diversity. Defining regeneration recruitment composition as being proportional to local adult abundance, regardless of productivity, might be a simple and conservative assumption to safely maintain relative species abundances (Hanbury-Brown, Ward, and Kueppers 2022), although this may be simplistic. Furthermore, based on the simulation results, there is no evidence that models with feedback from the canopy (in terms of species composition of regeneration) captured better the species diversity of the regeneration recruitment better than those without feedback, with the exception of ForCEEPS that featured significantly improved regeneration more accurate recruitment species diversity with the model version that included feedback (ForCEEPS(f)). The similar performance of models with and without feedback is likely because the models put more weight on the regeneration niche arising from abiotic and biotic filters, than from the habitat niche of the adult trees (Grubb 1977).

Regeneration Recruitment mortality

There are several factors that lead to mortality during the regeneration phase, such as competition (Casper and Jackson 1997) and multiple abiotic factors (Schmid et al. 2021; Cunningham et al. 2006). As mentioned above, it is reasonable to expect that models that overestimate regeneration recruitment at 7 cm may have a particularly high mortality between the 7 and 10 cm thresholds. However, this was not consistently evident from the simulations. This implies that these models must have an excess mortality in later stages, if we assume that all models were able to capture the structure and composition of the adult stands along the EuFoRIa gradient; this however was not tested here.

Higher mortality towards the adult phase has important implications for forest dynamics and the goods and services provided by forests. On the one hand, mortality in later stages may erroneously enhance the share of less shade tolerant species in the models (Klopčič, Simončič, and Bončina 2015), therefore shifting the species composition. Unrealistic high stem densities for a longer period of time may overestimate the role of tree regrowth in carbon sink dynamics (Pugh et al. 2019). Given our set of simulations and analyses, we cannot conclusively assess what is happening in the models, and further studies are required that focus on a wider range of tree sizes and the fate of tree regeneration along such a size continuum.

Model performance and model traits

In spite of the critical considerations above, it is remarkable that most models did not deviate exceedingly from the observations with respect to simulated regeneration recruited levels – after all, few if any of the models' regeneration routines are sufficiently well constrained by data, with the exception of the empirical models xComp and SIBYLA. It is noteworthy that in spite of their empirical basis, these two models did not match empirical regeneration recruitment levels, in a similar magnitude as the other models and even in

opposite directions (SIBYLA: underestimation; xComp: overestimation). It appears that using empirical data of limited geographical scope and that are constrained to managed forests, as done in these models, leads to extrapolation problems already under current climatic conditions. For the other models, whose regeneration modules are not strongly constrained by empirical data, the degrees of freedom multiple strategies that are available for modeling regeneration processes (König et al. 2022) could have implied that model performance would be much worse than what we found.

Our study showed that increasing complexity in the regeneration modules is not linked with to a higher accuracy of the projections of regeneration recruitment levels, species composition or mortality at early tree stages, as there was no significant relationship between model performance and model complexity. Increasing complexity in regeneration modules has been called for a long time ago (e.g., Price et al. (2001)) motivated by better process understanding and with the goal of enhancing of model accuracy (Bugmann and Seidl 2022). However, more complex models do may not necessarily lead to better projections and but rather to a reduced transparency and lower predictive power (Franklin et al. 2020). Thus, the question of what regarding the level of detail that is appropriate and parsimonious for modeling tree regeneration processes remains open (König et al. 2022; Bugmann and Seidl 2022).

Regeneration gradients on regeneration niches

Competition for light as a strong filter for tree regeneration has been widely documented (Messier et al. 1999; Collet and Chenost 2006; Berdanier and Clark 2016), but the models examined here did not reproduce this expectation. (i.e., decreasing recruitment levels with increasing stand basal area). However, it is difficult to measure light availability at large spatial and temporal scales. We used total stand basal area as a proxy for light availability (cf. Schmid et al. 2021). However Yet, we were unable to consider light availability restrictions caused by ground vegetation, which may be an important filter for some EuForIa forest dynamics at least in some EuForIa stands (Woltjer et al. 2008). We found pronounced differences in the ranges of stand basal area ranges simulated by the models, but also between models and observations. This made it impossible to evaluate the regeneration tree recruitment for the extremes of the stand density ranges in some models. For example, regeneration This is unfortunate because recruitment levels at low stand densities density are relevant to assess how well forests are recovering e.g. after gap creation due to disturbance (Seidl and Turner 2022; Grubb 1977). At the level of the simulated one-hectare samples, average basal area is typically not very low as long as no larger disturbances occur, which was explicitly excluded in our simulation protocol. Towards the other end of the spectrum, i.e., with increasing stand basal area, it would be reasonable to expect that the regeneration recruitment of the different species would become sparser and drop out entirely under low-light conditions (Klopčič, Poljanec, and Boncina 2012; Zell et al. 2019). However, few models showed this trend, thus indicating that the relation between regeneration and light availability is not yet captured correctly in most models. Yet, several of the models that did feature an increase of regeneration with increasing basal area include a feedback between seed production of mature trees and regeneration. Thus, it seems that in these models higher light competition does not

sufficiently compensate for increased seed availability with higher basal area due to a higher abundance of mature trees.

~~There~~Lastly, there were pronounced differences how the main tree species were represented by the models along the environmental gradients, in particular the dominant species *Abies alba*, *Fagus sylvatica* and *Picea abies*. The regenerationrecruitment levels were sampled from simulations in the equilibrium, and in this state it is expected that non-climax species such as *Pinus sylvestris* or *Quercus* would be of minor importance, or absent (Klopčič, Simončič, and Bončina 2015). Most models captured this low abundance, which is also found in the empirical data. Thus, although the broad patterns are matched by many models, improvements in the quantification of the regeneration niche of the species are needed, but this cannot be done in the absence of robust datasets across multiple species.

Methodological considerations

The EuFoRia data as used here are unique and well suited to better understand tree regeneration. However, three aspects of these data may represent considerable limitations. First, we made a comparison of tree regeneration in an equilibrium state, but we cannot assess how close the forests included in the EuFoRia dataset were to such a state. The data were collected in forest reserves where no management has taken place for long periods of time. This makes our assumption of an equilibrium between forest properties and environmental drivers more reasonable than it might appear at first sight. In an analysis of primeval forests at demographic equilibrium, Brzeziecki et al. (2021) found higher regenerationrecruitment rates than the ones observed in the EuFoRia data. Thus, the overestimation of regenerationrecruitment rates by the models may not be so problematic. Second, the data were collected for regenerationrecruitment above a 7 cm threshold, thus limiting the assessment of tree regeneration to a holistic viewpointspecific point of stand dynamics. This constituted a hard limit based on which we can understand only some aspects of tree regeneration, which in its entirety often comprises a rather long period since seed production (Price et al. 2001). In reality, many environmental constraints are acting on young trees (Käber et al. 2021) that we were unable to assess. Yet, regenerationrecruitment data with lower thresholds are simply not available in a harmonized manner across large environmental gradients. Third, the empirical data were collected from rather small plots, while we sampled simulated regenerationrecruitment levels from 1 ha areas, which may lead to an incorrect representation of space. Even though the strategy we adopted is not ideal, it represents a common challenge when harmonizing diverse data sources originating from varied sampling strategies (Portier et al. 2022). It would have been extremely challenging for such a variety of models to follow a protocol where the spatial sampling size was different at each of the 200 sites, and it would have introduced additional uncertainty in the results. ~~Ultimately, we need to acknowledge that the observed data represent just a snapshot of an extremely complex and stochastic reality. This may at least partly explain e.g. the lack of clear patterns in the simulation results along the gradients of climate water balance and degree-days.~~

The design of our sampling protocol did not include spatial aspects such as seed dispersal or detailed soil data. While we considered a wide range of models of forest dynamics, from stand to global scales, the simulation setup was limited to a spatial scale of 1 ha

regeneration values. This lack of consideration of spatial scale **relationships** is appropriate for stand-scale models, but it potentially puts the landscape models at a disadvantage, as they have been built to be accurate at the landscape level. **Without:** **without any** spatial context, we are limiting **tree** regeneration to the seed influx from the stand itself, unless the model has a background seed input. Yet, the global models should not be at a disadvantage **here** due to **the****their** inherently limited spatial considerations (but cf. Snell et al. 2018; or Lehsten et al. 2019), as they are usually lacking dispersal between cells and are based on a strong abstraction of horizontal space (Hanbury-Brown, Ward, and Kueppers 2022). Lastly, detailed data on soil conditions were not available from the observed data, and independent, admittedly **rather** coarse data for soil **properties** and **the** climatic water balance had to be used instead. It is noteworthy that many models represent drought **based** **on****using** detailed indicators based e.g. on soil water holding capacity, which had to be derived from a rough soil quality measure. This may at least partially explain the unsatisfactory performance of many models along the drought axis (i.e., climatic water balance).

Research recommendations

With our study, we have demonstrated that models of forest dynamics need a focus on **tree****their** regeneration modules to make them more robust. It remains uncertain what level of detail is required to model **tree** regeneration **processes**, and this must be addressed in future research. We recommend that the improvement of the regeneration modules is implemented as additional features that can be traced back, as done here for the variants of ForClim and ForCEEPS, and that model complexity and structure must always be connected with modeling objectives (Albrich et al. 2020). If it should be necessary to include more detail in the regeneration models, this will come with higher parameterisation efforts. This will most likely lead to lower generalization because the required data will have to be collected from specific locations, as currently there is no general, comprehensive regeneration dataset available.

Therefore, we further recommend that more effort should be invested into collecting harmonized datasets **on****in a site-specific manner covering the different aspects leading to** tree regeneration. We emphasize that datasets such as EuFoRia are invaluable and should be expanded in both their spatial extent (e.g., towards boreal and Mediterranean conditions) as well as in time (e.g., continuing the monitoring into the future). Such data will allow for a better evaluation of forest models and help to reduce the uncertainty in their projections, which is crucial when they are used as tools for predicting e.g. the impacts of anthropogenic climate change.

In the present study, we have considered tree regeneration in the equilibrium state only. It is equally important to understand how these models project tree regeneration after changes in forest structure by disturbances (Seidl and Turner 2022), or under different management strategies (Lindner, Lasch, and Erhard 2000). However, this will require an entirely different set of observed data, and potentially not all **models of** forest **dynamics****models** would be able to assess the relationship of these aspects on tree regeneration, e.g. due to the lack of disturbance or appropriate management modules.

We recommend to investigate in detail the implications of the current modeling strategies for tree regeneration and, ultimately, simulated forest stand structure. This applies particularly to the erroneous patterns of excess tree regeneration and later excess mortality, by focusing on a wider range of tree sizes and the related regeneration dynamics. We also recommend, especially for the landscape-level models, the inclusion of explicit spatial considerations regarding tree regeneration (Beckage and Clark 2003); this, however, is a serious challenge regarding appropriate datasets, as large-scale inventory data have a wide coverage, but by definition do not allow for the assessment of spatial interaction-effectsinteractions.

Exercises like the one presented here, where the models are operated in “blind flight” mode, i.e., without the possibility of being tuned towards capturing the expected patterns, should be repeated. Such benchmarking exercises should next focus on aspects such as specific model traits and the ecological formulations of particular (sub-)processes, to better understand the implications of the assumptions on which the models are based. Furthermore, the inclusion of a wide range of models with different (1) scales, type,(2) approaches for capturing population structures,structure,(3) tree regeneration modules and (4) complexity of their formulations will ensure a widerlarge benefit to the wholeentire modeling community and beyond.

Conclusions

Models of forest dynamics are important tools in science and decision support, and the formulation of tree regeneration has strong implications for simulated forest properties. The 15 models and variants used here are facing similar challenges in their representation of tree regeneration: they generally overestimate tree regenerationrecruitment levels, and the simulated regeneration niche is not always captured accurately as a function of biotic (light) and abiotic (temperature and moisture) factors.

However, most models properly capture the diversity of the initial tree community, and differences between model formulations, e.g., the presence or absence of feedback from the adult trees, did not have a strong effect foron capturing the species composition of regeneration.

Regarding mortality in the early phase of tree life, many models that feature a particularly high overestimation of regenerationrecruitment levels are compensating for this by a larger tree mortality. Often, this compensation is not sufficient to reduce the high regenerationrecruitment levels to realistic values. Overall, there is no clear mortality pattern across all models.

When capturing tree regeneration, the specific design decisions taken in the development of any model are more important for its behavior (accuracy) than scale (stand, landscape, global), modelling approach (empirical vs. process-based), and complexity. Having both empirical and process-based models in our set, the empirically-based models could have been expected to have a better performance, as they were calibrated with inventory data, but this was not the case. Similarly, higher model complexity does not represent an improvement for capturing tree regeneration.

Even though the regeneration routines of most of the models investigated here have never been sufficiently constrained well by robust data, the regeneration their projections of recruitment are not overly off. This indicates that a lot can be gained by a focus on the modeling of regeneration processes. The representation of forest dynamics in these models would become much more robust particularly in the face of climate change and post-disturbance dynamics, thus strongly reducing the uncertainty in long-term projections of future forest dynamics.

Author contributions

~~O. Díaz-Yáñez and Y. Käber contributed equally and share the first authorship. H. Bugmann, Y. Käber, and O. Díaz-Yáñez developed the study design. All authors contributed to preparing the simulations with their respective models. O. Díaz-Yáñez led the data analysis with contributions during the workshop from Y. Käber, T. Anders, K. H. Bražiūnas, J. Brůna, S. M. Fischer, J. Hetzer, T. Hickler, H. Lischke, M. Mahnken, P. Mairoti, K. Merganičová, T. Mette, X. Morin, W. Rammer and D. Scherrer. H. Bugmann, O. Díaz-Yáñez and H. Bugmann led the writing of the manuscript. All authors participated in the revision of the manuscript and approved its submission.~~

Acknowledgements

This paper arises from a workshop held in Davos (Switzerland) in June 2022, co-led by Olalla Díaz-Yáñez, Yannek Käber and Harald Bugmann. Funding for the workshop was provided by COST Action CA19139 PROCLIAS (PROcess-based models for CLimate Impact Attribution across Sectors), supported by COST (European Cooperation in Science and Technology; <https://www.cost.eu>) and ETH Zurich (Research Grant ETH-35 18-1). ODY was funded by the Swiss National Science Foundation project 200020_188882 / 1. KB received support from the European Research Council under the European Union's Horizon 2020 research and innovation program (Grant Agreement 101001905). JB acknowledges funding from the long-term research development project no. RVO 67985939 from the Czech Academy of Sciences. MM acknowledges financial support from I-Maestro (Innovative forest management strategies for a resilient bioeconomy under climate change and disturbances, grant no. 22035418, 2019–2022) funded by the ERA-NET Cofund ForestValue. MMina acknowledges funding from the H2020-MSCA-IF project REINFORCE (Grant No. 891671). KM received funding from the Project No. CZ.02.1.01/0.0/0.0/16_019/0000803 financed by OP RDE, grant no. ITMS2014+313011W580 supported by the Integrated Infrastructure Operational Programme funded by the ERDF. CH acknowledges financial support from Waldfonds WASIM2100 project. We would also like to thank all the people thatwho helped to make this study possible with contributions at different stages such asparticularly Martin Gutsch and Giorgio Vacchiano.

Author contributions

O. Díaz-Yáñez and Y. Käber contributed equally and share the first authorship. H. Bugmann, Y. Käber, and O. Díaz-Yáñez developed the study design. Y. Käber and O. Díaz-Yáñez prepared the input data used by all modeling teams. All authors contributed by conducting the simulations with their respective models. O. Díaz-Yáñez led the analysis of the simulation results with contributions during the workshop from Y. Käber, T. Anders, K. H. Braziunas, J. Brůna, S. M. Fischer, J. Hetzer, T. Hickler, H. Lischke, M. Mahnken, P. Mairoti, K. Merganičová, T. Mette, X. Morin, W. Rammer, and D. Scherrer. O. Díaz-Yáñez and H. Bugmann led the writing of the manuscript and provided the first draft. All authors participated in the revision of the manuscript and approved its submission.

Conflict of interest statement

The authors declare no conflicts of interest.

References

- Albrich, Katharina, Werner Rammer, Monica G. Turner, Zak Ratajczak, Kristin H. Braziunas, Winslow D. Hansen, and Rupert Seidl. 2020. "Simulating Forest Resilience: A Review." Edited by Thomas Hickler. *Global Ecology and Biogeography* 29 (12): 2082–96. <https://doi.org/10.1111/geb.13197>.
- Allen, Jon C. 1976. "A Modified Sine Wave Method for Calculating Degree Days1." *Environmental Entomology* 5 (3): 388–96. <https://doi.org/10.1093/ee/5.3.388>.
- ASTER Science Team. 2019. "ASTER Global Digital Elevation Model V003. 2019, Distributed by NASA EOSDIS Land Processes DAAC, <Https://Doi.org/10.5067/ASTER/ASTGTM.003>. Accessed 2022-11-18." <Https://lpdaac.usgs.gov/products/astgtmv002/>.
- Beckage, Brian, and James S. Clark. 2003. "SEEDLING SURVIVAL AND GROWTH OF THREE FOREST TREE SPECIES: THE ROLE OF SPATIAL HETEROGENEITY." *Ecology* 84 (7): 1849–61. [Https://doi.org/10.1890/0012-9658\(2003\)084\[1849:SSAGOT\]2.0.CO;2](Https://doi.org/10.1890/0012-9658(2003)084[1849:SSAGOT]2.0.CO;2).
- Berdanier, Aaron B., and James S. Clark. 2016. "Divergent Reproductive Allocation Trade-Offs with Canopy Exposure Across Tree Species in Temperate Forests." *Ecosphere* 7 (6): e01313. <Https://doi.org/10.1002/ecs2.1313>.
- Bogdziewicz, Michał, Andrew Hacket-Pain, Dave Kelly, Peter A. Thomas, Jonathan Lageard, and Andrew J. Tanentzap. 2021. "Climate Warming Causes Mast Seeding to Break down by Reducing Sensitivity to Weather Cues." *Global Change Biology* 27 (9): 1952–61. <Https://doi.org/10.1111/gcb.15560>.
- Bohn, Friedrich J., Karin Frank, and Andreas Huth. 2014. "Of Climate and Its Resulting Tree Growth: Simulating the Productivity of Temperate Forests." *Ecological Modelling* 278 (April): 9–17. <Https://doi.org/10.1016/j.ecolmodel.2014.01.021>.

Botkin, Daniel B., James F. Janak, and James R. Wallis. 1972. "Some Ecological Consequences of a Computer Model of Forest Growth." *The Journal of Ecology* 60 (3): 849. <https://doi.org/10.2307/2258570>.

Brzeziecki, Bogdan, Stanisław Drozdowski, Kamil Bielak, Marcin Czacharowski, Jacek Zajączkowski, Włodzimierz Buraczyk, and Leszek Gawron. 2021. "A Demographic Equilibrium Approach to Stocking Control in Mixed, Multiaged Stands in Białowieża Forest, Northeast Poland." *Forest Ecology and Management* 481 (February): 118694. <https://doi.org/10.1016/j.foreco.2020.118694>.

Bugmann, Harald. 1996. "A Simplified Forest Model to Study Species Composition Along Climate Gradients." *Ecology* 77 (7): 2055–74. <https://doi.org/10.2307/2265700>.

Bugmann, Harald, and Rupert Seidl. 2022. "The Evolution, Complexity and Diversity of Models of Long-Term Forest Dynamics." *Journal of Ecology* 110 (10): 2288–2307. <https://doi.org/10.1111/1365-2745.13989>.

Bugmann, Harald, Rupert Seidl, Florian Hartig, Friedrich Bohn, Josef Brůna, Maxime Cailleret, Louis François, et al. 2019. "Tree Mortality Submodels Drive Simulated Long-Term Forest Dynamics: Assessing 15 Models from the Stand to Global Scale." *Ecosphere* 10 (2): e02616. <https://doi.org/10.1002/ecs2.2616>.

Bugmann, Harald, Xiaodong Yan, MartinT. Sykes, Philippe Martin, Marcus Lindner, PaulV. Desanker, and SteveG. Cumming. 1996. "A Comparison of Forest Gap Models: Model Structure and Behaviour." *Climatic Change* 34 (2). <https://doi.org/10.1007/BF00224640>.

Cailleret, Maxime, Steven Jansen, Elisabeth M. R. Robert, Lucía Desoto, Tuomas Aakala, Joseph A. Antos, Barbara Beikircher, et al. 2017. "A Synthesis of Radial Growth Patterns Preceding Tree Mortality." *Global Change Biology* 23 (4): 1675–90. <https://doi.org/10.1111/gcb.13535>.

Casper, Brenda B., and Robert B. Jackson. 1997. "Plant Competition Underground." *Annual Review of Ecology and Systematics* 28 (1): 545–70. <https://doi.org/10.1146/annurev.ecolsys.28.1.545>.

Clark, J. S., B. Beckage, P. Camill, B. Cleveland, J. HilleRisLambers, J. Lichter, J. McLachlan, J. Mohan, and P. Wyckoff. 1999. "Interpreting Recruitment Limitation in Forests." *American Journal of Botany* 86 (1): 1–16. <https://doi.org/10.2307/2656950>.

Collet, C., and C. Chenost. 2006. "Using Competition and Light Estimates to Predict Diameter and Height Growth of Naturally Regenerated Beech Seedlings Growing Under Changing Canopy Conditions." *Forestry* 79 (5): 489–502. <https://doi.org/10.1093/forestry/cpl033>.

Cunningham, Catherine, Niklaus E Zimmermann, Veronika Stoeckli, and Harald Bugmann. 2006. "Growth Response of Norway Spruce Saplings in Two Forest Gaps in the Swiss Alps to Artificial Browsing, Infection with Black Snow Mold, and Competition by Ground Vegetation." *Canadian Journal of Forest Research* 36 (11): 2782–93. <https://doi.org/10.1139/x06-156>.

- Díaz-Yáñez, Olalla, Yannek Käber, and Harald Bugmann. 2022. "Regeneration Workshop Protocol." OSF. <https://doi.org/10.17605/OSF.IO/CZDXP>.
- EuFoRia. 2019. "European Forest Reserves Initiative. Link: <Https://Www.wsl.ch/de/Projekte/Euforia.html#tabelement1-Tab2>."
- Fabrika, M. 2005. "Simulátor Biodynamiky Lesa SIBYLA, Koncepcia, Konštrukcia a Programové Riešenie." *Zvolen, Slovakia, Habilitačná Práca, Technická Univerzita Vo Zvolene.*
- Fabrika, M., and J. Ďurský. 2012. "Algorithms and Software Solution of Thinning Models for SIBYLA Growth Simulator." *Journal of Forest Science* 51 (No. 10): 431–45. <https://doi.org/10.17221/4577-JFS>.
- Fischer, Rico, Friedrich Bohn, Mateus Dantas de Paula, Claudia Dislich, Jürgen Groeneveld, Alvaro G. Gutiérrez, Martin Kazmierczak, et al. 2016. "Lessons Learned from Applying a Forest Gap Model to Understand Ecosystem and Carbon Dynamics of Complex Tropical Forests." *Ecological Modelling* 326 (April): 124–33. <https://doi.org/10.1016/j.ecolmodel.2015.11.018>.
- Fischlin, Andreas, Harald Bugmann, and Dimitrios Gyalistras. 1995. "Sensitivity of a Forest Ecosystem Model to Climate Parametrization Schemes." *Environmental Pollution* 87 (3): 267–82. [https://doi.org/10.1016/0269-7491\(94\)P4158-K](https://doi.org/10.1016/0269-7491(94)P4158-K).
- FOREST EUROPE. 2020. "State of Europe's Forests 2020." Ministerial Conference on the Protection of Forests in Europe - FOREST EUROPE.
- Fortin, Mathieu, and Josianne DeBlois. 2007. "Modeling Tree Recruitment with Zero-Inflated Models: The Example of Hardwood Stands in Southern Québec, Canada." *Forest Sciences* 53 (4): 529–39. <https://doi.org/10.1093/forestscience/53.4.529>.
- Franklin, Oskar, Sandy P. Harrison, Roderick Dewar, Caroline E. Farrior, Åke Brännström, Ulf Dieckmann, Stephan Pietsch, et al. 2020. "Organizing Principles for Vegetation Dynamics." *Nature Plants* 6 (5): 444–53. <https://doi.org/10.1038/s41477-020-0655-x>.
- Grubb, P. J. 1977. "THE MAINTENANCE OF SPECIES-RICHNESS IN PLANT COMMUNITIES: THE IMPORTANCE OF THE REGENERATION NICHE." *Biological Reviews* 52 (1): 107–45. <https://doi.org/10.1111/j.1469-185X.1977.tb01347.x>.
- Gutiérrez, Alvaro G., Rebecca S. Snell, and Harald Bugmann. 2016. "Using a Dynamic Forest Model to Predict Tree Species Distributions: Predicting Tree Species Distribution." *Global Ecology and Biogeography* 25 (3): 347–58. <https://doi.org/10.1111/geb.12421>.
- Hallsby, Göran, Kristina Ahnlund Ulvcrona, Anders Karlsson, Björn Elfving, Hans Sjögren, Thomas Ulvcrona, and Urban Bergsten. 2015. "Effects of Intensity of Forest Regeneration Measures on Stand Development in a Nationwide Swedish Field Experiment." *Forestry* 88 (4): 441–53. <https://doi.org/10.1093/forestry/cpv010>.

- Hanbury-Brown, Adam R., Rachel E. Ward, and Lara M. Kueppers. 2022. "Forest Regeneration Within Earth System Models: Current Process Representations and Ways Forward." *New Phytologist* 235 (1): 20–40. <https://doi.org/10.1111/nph.18131>.
- Heiri, C., A. Wolf, L. Rohrer, and H. Bugmann. 2009. "Forty Years of Natural Dynamics in Swiss Beech Forests: Structure, Composition, and the Influence of Former Management." *Ecological Applications* 19 (7): 1920–34. <https://doi.org/10.1890/08-0516.1>.
- Hengl, Tomislav, Jorge Mendes de Jesus, Gerard B. M. Heuvelink, Maria Ruiperez Gonzalez, Milan Kilibarda, Aleksandar Blagotić, Wei Shangguan, et al. 2017. "SoilGrids250m: Global Gridded Soil Information Based on Machine Learning." Edited by Ben Bond-Lamberty. *PLOS ONE* 12 (2): e0169748. <https://doi.org/10.1371/journal.pone.0169748>.
- Hickler, Thomas, Katrin Vohland, Jane Feehan, Paul A. Miller, Benjamin Smith, Luis Costa, Thomas Giesecke, et al. 2012. "Projecting the Future Distribution of European Potential Natural Vegetation Zones with a Generalized, Tree Species-Based Dynamic Vegetation Model: Future Changes in European Vegetation Zones." *Global Ecology and Biogeography* 21 (1): 50–63. <https://doi.org/10.1111/j.1466-8238.2010.00613.x>.
- Huber, Nica, Harald Bugmann, Maxime Cailleret, Nicolas Bircher, and Valentine Lafond. 2021. "Stand-Scale Climate Change Impacts on Forests over Large Areas: Transient Responses and Projection Uncertainties." *Ecological Applications* 31 (4). <https://doi.org/10.1002/eap.2313>.
- Huber, Nica, Harald Bugmann, and Valentine Lafond. 2020. "Capturing Ecological Processes in Dynamic Forest Models: Why There Is No Silver Bullet to Cope with Complexity." *Ecosphere* 11 (5). <https://doi.org/10.1002/ecs2.3109>.
- Käber, Yannek, Christof Bigler, ..., and Harald Bugmann. 2023. "Sheltered and Suppressed: Tree Regeneration in Unmanaged European Forests (Under Revision)." *Journal of Ecology*, December.
- Käber, Yannek, Peter Meyer, Jonas Stillhard, Emiel De Lombaerde, Jürgen Zell, Golo Stadelmann, Harald Bugmann, and Christof Bigler. 2021. "Tree Recruitment Is Determined by Stand Structure and Shade Tolerance with Uncertain Role of Climate and Water Relations." *Ecology and Evolution* 11 (17): 12182–203. <https://doi.org/10.1002/ece3.7984>.
- Karger, Dirk N., Stefan Lange, Chantal Hari, Christopher P. O. Reyer, and Niklaus E. Zimmermann. 2021. "CHELSA-W5E5 V1.0: W5E5 V1.0 Downscaled with CHELSA V2.0. ISIMIP Repository. [Https://Doi.org/10.48364/ISIMIP.836809](https://doi.org/10.48364/ISIMIP.836809)."
- Klopčič, Matija, Ales Poljanec, and Andrej Boncina. 2012. "Modelling Natural Recruitment of European Beech (*Fagus Sylvatica L.*)."*Forest Ecology and Management* 284 (November): 142–51. <https://doi.org/10.1016/j.foreco.2012.07.049>.
- Klopčič, Matija, Tina Simončič, and Andrej Bončina. 2015. "Comparison of Regeneration and Recruitment of Shade-Tolerant and Light-Demanding Tree Species in Mixed Uneven-Aged

Forests: Experiences from the Dinaric Region." *Forestry* 88 (5): 552–63. <https://doi.org/10.1093/forestry/cpv021>.

König, Louis A., Frits Mohren, Mart-Jan Schelhaas, Harald Bugmann, and Gert-Jan Nabuurs. 2022. "Tree Regeneration in Models of Forest Dynamics – Suitability to Assess Climate Change Impacts on European Forests." *Forest Ecology and Management* 520 (September): 120390. <https://doi.org/10.1016/j.foreco.2022.120390>.

Kroiss, Steve J., and Janneke HilleRisLambers. 2015. "Recruitment Limitation of Long-Lived Conifers: Implications for Climate Change Responses." *Ecology* 96 (5): 1286–97. <https://doi.org/10.1890/14-0595.1>.

Langan, Liam, Steven I. Higgins, and Simon Scheiter. 2017. "Climate-Biomes, Pedo-Biomes or Pyro-Biomes: Which World View Explains the Tropical Forest-Savanna Boundary in South America?" *Journal of Biogeography* 44 (10): 2319–30. <https://doi.org/10.1111/jbi.13018>.

Lasch-Born, Petra, Felicitas Suckow, Christopher P. O. Reyer, Martin Gutsch, Chris Kollas, Franz-Werner Badeck, Harald K. M. Bugmann, et al. 2020. "Description and Evaluation of the Process-Based Forest Model 4C V2.2 at Four European Forest Sites." *Geoscientific Model Development* 13 (11): 5311–43. <https://doi.org/10.5194/gmd-13-5311-2020>.

Lehsten, Veiko, Michael Mischirow, Erik Lindström, Dörte Lehsten, and Heike Lischke. 2019. "LPJ-GM 1.0: Simulating Migration Efficiently in a Dynamic Vegetation Model." *Geoscientific Model Development* 12 (3): 893–908. <https://doi.org/10.5194/gmd-12-893-2019>.

Leishman, Mr, L Hughes, K French, D Armstrong, and M Westoby. 1992. "Seed and Seedling Biology in Relation to Modelling Vegetation Dynamics Under Global Climate Change." *Australian Journal of Botany* 40 (5): 599. <https://doi.org/10.1071/BT9920599>.

Lexer, Manfred J, and Karl Hönniger. 2001. "A Modified 3D-Patch Model for Spatially Explicit Simulation of Vegetation Composition in Heterogeneous Landscapes." *Forest Ecology and Management*, 23.

Lindner, Marcus, Petra Lasch, and Markus Erhard. 2000. "Alternative Forest Management Strategies Under Climatic Change Prospects for Gap Model Applications in Risk Analyses." *Silva Fennica* 34 (2). <https://doi.org/10.14214/sf.634>.

Lindner, Marcus, Michael Maroschek, Sigrid Netherer, Antoine Kremer, Anna Barbati, Jordi Garcia-Gonzalo, Rupert Seidl, et al. 2010. "Climate Change Impacts, Adaptive Capacity, and Vulnerability of European Forest Ecosystems." *Forest Ecology and Management* 259 (4): 698–709. <https://doi.org/10.1016/j.foreco.2009.09.023>.

Lischke, Heike, and Thomas J. Löffler. 2006. "Intra-Specific Density Dependence Is Required to Maintain Species Diversity in Spatio-Temporal Forest Simulations with Reproduction." *Ecological Modelling* 198 (3-4): 341–61. <https://doi.org/10.1016/j.ecolmodel.2006.05.005>.

Meier, Eliane S., Thomas C. Edwards Jr, Felix Kienast, Matthias Dobbertin, and Niklaus E. Zimmermann. 2011. "Co-Occurrence Patterns of Trees Along Macro-Climatic Gradients and Their Potential Influence on the Present and Future Distribution of *Fagus Sylvatica* L.: Influence of Co-Occurrence Patterns on *Fagus Sylvatica*." *Journal of Biogeography* 38 (2): 371–82. [https://doi.org/10.1111/j.1365-2699.2010.02405.x.](https://doi.org/10.1111/j.1365-2699.2010.02405.x)

Messier, Christian, René Doucet, Jean-Claude Ruel, Yves Claveau, Colin Kelly, and Martin J Lechowicz. 1999. "Functional Ecology of Advance Regeneration in Relation to Light in Boreal Forests" 29: 12.

Mette, Tobias. 2014. "Modelling Patagonian Lenga-Forest Dynamics (*Nothofagus Pumilio*) in Chile. Final Report, DFG Project ME 3568/2-1, Bonn, Germany."

Miina, Jari, Kalle Eerikäinen, and Hubert Hasenauer. 2006. "Modeling Forest Regeneration. In: Hasenauer, h. (Eds.), Sustainable Forest Management. Growth Models for Europe (Pp. 93–110)." In.

Morin, Xavier, Harald Bugmann, François Coligny, Nicolas Martin-StPaul, Maxime Cailleret, Jean-Marc Limousin, Jean-Marc Ourcival, et al. 2021. "Beyond Forest Succession: A Gap Model to Study Ecosystem Functioning and Tree Community Composition Under Climate Change." Edited by Julia Koricheva. *Functional Ecology* 35 (4): 955–75.
<https://doi.org/10.1111/1365-2435.13760>.

Muñoz Sabater, J. 2019. "ERA5-Land Hourly Data from 1981 to Present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). (23.01.2022)." [doi:10.24381/cds.e2161bac](https://doi.org/10.24381/cds.e2161bac).

Pan, Yude, Richard A. Birdsey, Jingyun Fang, Richard Houghton, Pekka E. Kauppi, Werner A. Kurz, Oliver L. Phillips, et al. 2011. "A Large and Persistent Carbon Sink in the World's Forests." *Science* 333 (6045): 988–93. <https://doi.org/10.1126/science.1201609>.

Portier, Jeanne, Florian Zellweger, Jürgen Zell, Iciar Alberdi Asensio, Michal Bosela, Johannes Breidenbach, Vladimír Šebeň, Rafael O. Wüest, and Brigitte Rohner. 2022. "Plot Size Matters: Toward Comparable Species Richness Estimates Across Plot-Based Inventories." *Ecology and Evolution* 12 (6). <https://doi.org/10.1002/ece3.8965>.

Pretzsch, Hans, and Peter Biber. 2003. "A Re-Evaluation of Reineke's Rule and Stand Density Index." *Forest Sciences* 51 (4): 304–20.
<https://doi.org/10.1093/forestscience/51.4.304>.

Pretzsch, H., P. Biber, and J. Ďurský. 2002. "The Single Tree-Based Stand Simulator SILVA: Construction, Application and Evaluation." *Forest Ecology and Management* 162 (1): 3–21.
[https://doi.org/10.1016/S0378-1127\(02\)00047-6](https://doi.org/10.1016/S0378-1127(02)00047-6).

Price, David T, Niklaus E Zimmermann, Manfred J Lexer, Paul Leadley, Irma T M Jorritsma, Jörg Schaber, Donald F Clark, et al. 2001. "Regeneration in Gap Models: Priority Issues for Studying Forest Responses to Climate Change" 51: 34.

Pugh, Thomas A. M., Mats Lindeskog, Benjamin Smith, Benjamin Poulter, Almut Arneth, Vanessa Haverd, and Leonardo Calle. 2019. "Role of Forest Regrowth in Global Carbon Sink

Dynamics." *Proceedings of the National Academy of Sciences* 116 (10): 4382–87. <https://doi.org/10.1073/pnas.1810512116>.

Rammer, Werner, Kristin H. Braziunas, Winslow D. Hansen, Zak Ratajczak, Anthony L. Westerling, Monica G. Turner, and Rupert Seidl. 2021. "Widespread Regeneration Failure in Forests of Greater Yellowstone Under Scenarios of Future Climate and Fire." *Global Change Biology* 27 (18): 4339–51. <https://doi.org/10.1111/gcb.15726>.

Redmond, Miranda D., Neil S. Cobb, Michael J. Clifford, and Nichole N. Barger. 2015. "Woodland Recovery Following Drought-Induced Tree Mortality Across an Environmental Stress Gradient." *Global Change Biology* 21 (10): 3685–95. <https://doi.org/10.1111/gcb.12976>.

Reineke, L H. 1933. "PERFECTING A STAND-DENSITY INDEX FOR EVEN- AGED FORESTS," 12.

Reyer, Christopher, Petra Lasch-Born, Felicitas Suckow, Martin Gutsch, Aline Murawski, and Tobias Pilz. 2014. "Projections of Regional Changes in Forest Net Primary Productivity for Different Tree Species in Europe Driven by Climate Change and Carbon Dioxide."

Ribbens, Eric, John A. Silander, and Stephen W. Pacala. 1994. "Seedling Recruitment in Forests: Calibrating Models to Predict Patterns of Tree Seedling Dispersion." *Ecology* 75 (6): 1794–1806. <https://doi.org/10.2307/1939638>.

Scheiter, Simon, and S. I. Higgins. 2009. "Impacts of Climate Change on the Vegetation of Africa: An Adaptive Dynamic Vegetation Modelling Approach." *Global Change Biology*.

Scheiter, Simon, Liam Langan, and Steven I. Higgins. 2013. "Next-Generation Dynamic Global Vegetation Models: Learning from Community Ecology." *New Phytologist* 198 (3): 957–69. <https://doi.org/10.1111/nph.12210>.

Scheller, Robert M., James B. Domingo, Brian R. Sturtevant, Jeremy S. Williams, Arnold Rudy, Eric J. Gustafson, and David J. Mladenoff. 2007. "Design, Development, and Application of LANDIS-II, a Spatial Landscape Simulation Model with Flexible Temporal and Spatial Resolution." *Ecological Modelling* 201 (3-4): 409–19. <https://doi.org/10.1016/j.ecolmodel.2006.10.009>.

Schmid, Ueli, Christof Bigler, Monika Frehner, and Harald Bugmann. 2021. "Abiotic and Biotic Determinants of Height Growth of Picea Abies Regeneration in Small Forest Gaps in the Swiss Alps." *Forest Ecology and Management* 490 (June): 119076. <https://doi.org/10.1016/j.foreco.2021.119076>.

Schumacher, Sabine, Björn Reineking, Jason Sibold, and Harald Bugmann. 2006. "Modeling the Impact of Climate and Vegetation on Fire Regimes in Mountain Landscapes." *Landscape Ecology* 21 (4): 539–54. <https://doi.org/10.1007/s10980-005-2165-7>.

Seidl, Rupert, Werner Rammer, Robert M. Scheller, and Thomas A. Spies. 2012. "An Individual-Based Process Model to Simulate Landscape-Scale Forest Ecosystem Dynamics."

Ecological Modelling 231 (April): 87–100.

<https://doi.org/10.1016/j.ecolmodel.2012.02.015>.

Seidl, Rupert, and Monica G. Turner. 2022. "Post-Disturbance Reorganization of Forest Ecosystems in a Changing World." *Proceedings of the National Academy of Sciences* 119 (28): e2202190119. <https://doi.org/10.1073/pnas.2202190119>.

Shugart, Herman Henry. 1984. *A Theory of Forest Dynamics. The Ecological Implications of Forest Succession Models*. NY, USA: Springer-Verlag.

Smith, Benjamin, I. Colin Prentice, and Martin T. Sykes. 2001a. "Representation of Vegetation Dynamics in the Modelling of Terrestrial Ecosystems: Comparing Two Contrasting Approaches Within European Climate Space: *Vegetation Dynamics in Ecosystem Models*." *Global Ecology and Biogeography* 10 (6): 621–37. <https://doi.org/10.1046/j.1466-822X.2001.t01-1-00256.x>.

———. 2001b. "Representation of Vegetation Dynamics in the Modelling of Terrestrial Ecosystems: Comparing Two Contrasting Approaches Within European Climate Space: *Vegetation Dynamics in Ecosystem Models*." *Global Ecology and Biogeography* 10 (6): 621–37. <https://doi.org/10.1046/j.1466-822X.2001.t01-1-00256.x>.

Snell, Rebecca S., Ché Elkin, Sven Kotlarski, and Harald Bugmann. 2018. "Importance of Climate Uncertainty for Projections of Forest Ecosystem Services." *Regional Environmental Change* 18 (7): 2145–59. <https://doi.org/10.1007/s10113-018-1337-3>.

Speich, Mjr. 2019. "Quantifying and Modeling Water Availability in Temperate Forests: A Review of Drought and Aridity Indices." *iForest - Biogeosciences and Forestry* 12 (1): 1–16. <https://doi.org/10.3832/ifor2934-011>.

Thriplleton, Timothy, Clemens Blattert, Leo Gallus Bont, Reinhard Mey, Jürgen Zell, Esther Thürig, and Janine Schweier. 2021. "A Multi-Criteria Decision Support System for Strategic Planning at the Swiss Forest Enterprise Level: Coping With Climate Change and Shifting Demands in Ecosystem Service Provisioning." *Frontiers in Forests and Global Change* 4 (August): 693020. <https://doi.org/10.3389/ffgc.2021.693020>.

Trugman, Anna T. 2022. "Integrating Plant Physiology and Community Ecology Across Scales Through Trait-Based Models to Predict Drought Mortality." *New Phytologist* 234 (1): 21–27. <https://doi.org/10.1111/nph.17821>.

Vacchiano, Giorgio, Davide Ascoli, Fabio Berzaghi, Manuel Esteban Lucas-Borja, Thomas Caignard, Alessio Collalti, Paola Mairotta, et al. 2018. "Reproducing Reproduction: How to Simulate Mast Seeding in Forest Models." *Ecological Modelling* 376 (May): 40–53. <https://doi.org/10.1016/j.ecolmodel.2018.03.004>.

Vanclay, J. K., and Jerome K. 1994. *Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests*. Wallingford, U.K: CAB International.

Vanclay, J. K., and J. P. Skovsgaard. 1997. "Evaluating Forest Growth Models." *Ecological Modelling* 98 (1): 1–12. [https://doi.org/10.1016/S0304-3800\(96\)01932-1](https://doi.org/10.1016/S0304-3800(96)01932-1).

Walck, Jeffrey L., Siti N. Hidayati, Kingsley W. Dixon, Ken Thompson, and Peter Poschlod. 2011. "Climate Change and Plant Regeneration from Seed." *Global Change Biology* 17 (6): 2145–61. <https://doi.org/10.1111/j.1365-2486.2010.02368.x>.

Wehrli, A., P. J. Weisberg, W. Schönenberger, P. Brang, and H. Bugmann. 2006. "Improving the Establishment Submodel of a Forest Patch Model to Assess the Long-Term Protective Effect of Mountain Forests." *European Journal of Forest Research* 126 (1): 131–45. <https://doi.org/10.1007/s10342-006-0142-6>.

Woltjer, M., W. Rammer, M. Brauner, R. Seidl, G. M. J. Mohren, and M. J. Lexer. 2008. "Coupling a 3D Patch Model and a Rockfall Module to Assess Rockfall Protection in Mountain Forests." *Journal of Environmental Management* 87 (3): 373–88. <https://doi.org/10.1016/j.jenvman.2007.01.031>.

Wood, S N. 2011. "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models." *Journal of the Royal Statistical Society* 73 (1): 3–36. <https://doi.org/doi.org/10.1111/j.1467-9868.2010.00749.x>.

Zell, Jürgen, Brigitte Rohner, Esther Thürig, and Golo Stadelmann. 2019. "Modeling Ingrowth for Empirical Forest Prediction Systems." *Forest Ecology and Management* 433 (February): 771–79. <https://doi.org/10.1016/j.foreco.2018.11.052>.

Table

Table 1. Models included in this study, their characteristics and simulation strategies used. Scale refers to the models scale, Type to the models type, Population structure to the model population approach, Feedback to the inclusion or not of adult tree feedback to the level and species composition of regeneration, Approach to the model regeneration module approach, Start from to the stage from which trees are recruited, Complexity, to the mean regeneration formulation as estimated in (Bugmann & Seidl, 2022), Species to which species from the original eleven species were simulated, Runtime (spin up) [years] to the number of years used per model in their simulations for the spin up period and regeneration sampling; Climate data type (use of years) to the type of data used in the simulations and how it was used, and Reference provides the link to the publication with more model details. Further information can be found on the individual model reports available in the supplementary materials.

For Review Only

Figure captions

Figure 1. The 200 sites of the EuFoRia network (EuFoRia, 2019) used in the present study, the colors show the dominant tree species per site in the observed regeneration at the 10 cm threshold.

Figure 2. Mean regeneration levels across all samples per site, plotted for the 200 sites and for each model. The red dashed lines shows the 25th and 75th percentiles for the 7 cm diameter threshold in the observed data. There are two boxplots for each model where the left lighter boxplot corresponds to 7 cm and the right darker boxplot to 10 cm

Figure 3. Shannon index of tree regeneration, calculated by basal area, as the mean value across the 200 samples per site for the diameter thresholds of 7 and 10 cm, respectively. The red dashed lines shows the 25th and 75th percentiles for the 7 cm diameter threshold in the observed data.

Figure 4. Mean Shannon index across all samples per site for observed and simulated data. Each plot shows one pattern represented by one exemplary model of each category (~~overpredicted overpredicted~~, intermediate and underpredicted). The full data with the grouping of the models are shown in Figures S3 and S4. n indicates the number of models falling in each group. A) Examples for the three trends across models for species diversity at the stand level. B) Examples for the three trends across models for regeneration at the 7 cm threshold. [\(165 sites\)](#).

Figure 5. Ratio of tree regeneration rates between the 7 and 10 cm thresholds. Dashed blue lines mark a ratio equal to 1, indicating no decrease in tree regeneration between 7 and 10 cm, and a ratio equal to 1.77, corresponding to the Reineke self-thinning ratio under even-aged conditions.

Figure 6. Ratio of tree regeneration between the 7 and 10 cm thresholds (regeneration 7 cm / regeneration 10 cm) and the overestimation proportion at 7 cm ([simulated - observed] / observed) for the mean regeneration per model across sites and samples. The horizontal dashed lines show a ratio equal to 1, indicating no decrease of regeneration between 7 and 10 cm, and a ratio equal to 1.77 corresponding to the Reineke self-thinning line under even-aged conditions. The vertical line indicates zero overestimation.

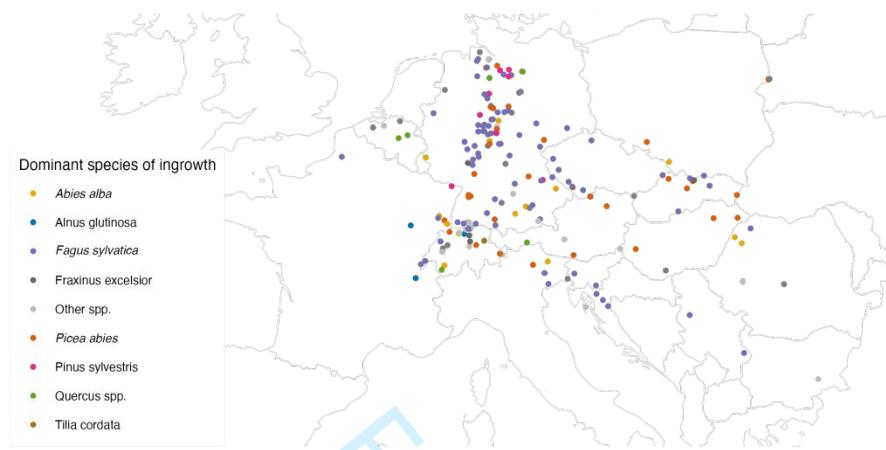
Figure 7. Mean ~~regeneration values recruitment~~ across the 200 samples per site with the y-axis scaled differently by model, for the 200 sites against gradients of (a) total basal area; (b) climatic water balance; (c) seasonal degree-day sum. The values were split into ten bins; the red lines represent a GAM model showing the trend in the observed data. ~~For comparing the magnitude of recruitment across models the same graphs with equally scaled y-axis can be found in Figure S7 Y-axis are not equally scaled.~~

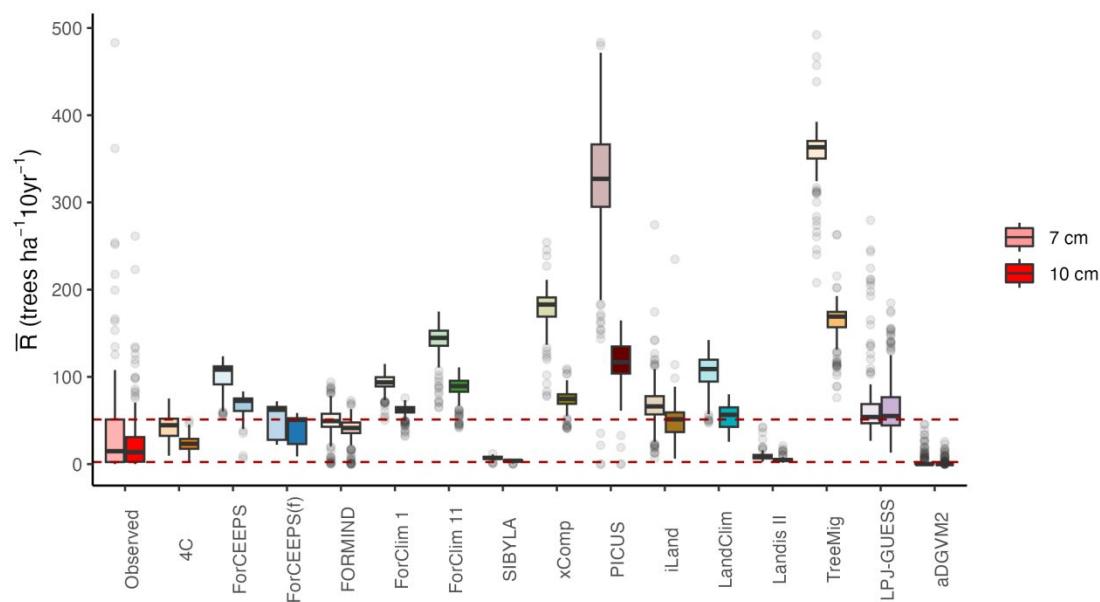
Figure 8. Regeneration levels of the five main species in terms of their basal area share in the observations (top row) and the performance of each model across the environmental gradients (other rows) for the 7 cm regeneration threshold. The values shown are the mean of the 200 samples per site and across the sites in each bin (tile), with ten bins per gradient. The sizes of the circles represent the ratio between the regeneration basal area of

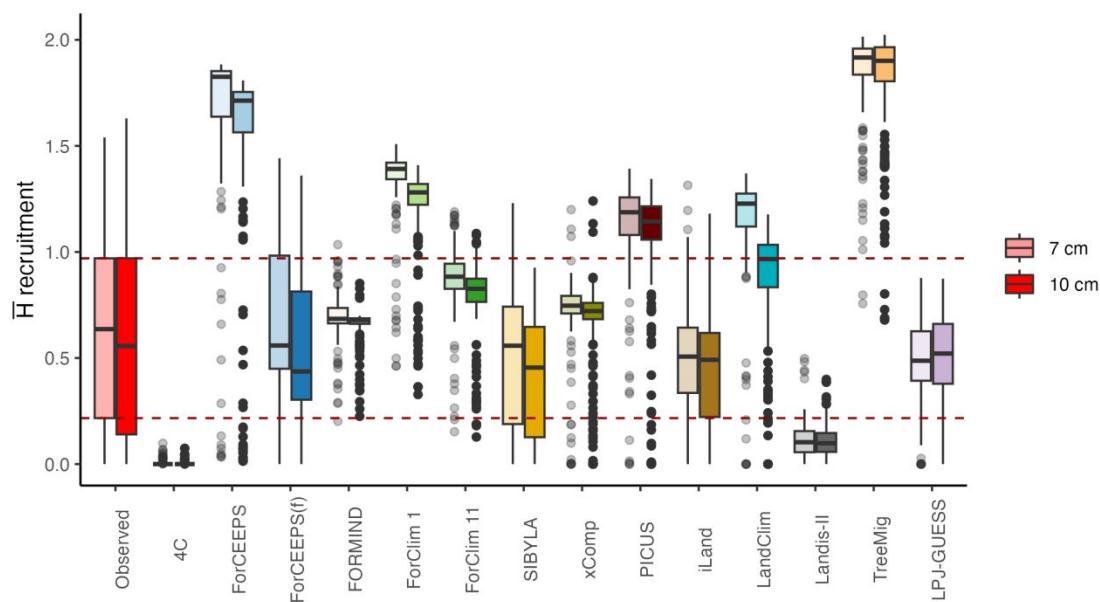
the species and the total regeneration basal area for all species. The absence of a circle indicates a zero basal area share in the regeneration, or the absence of regeneration altogether. The color gradient (for the models only) shows the difference between the simulated and observed ratio of regeneration basal area of the species and the total regeneration basal area for all the species.

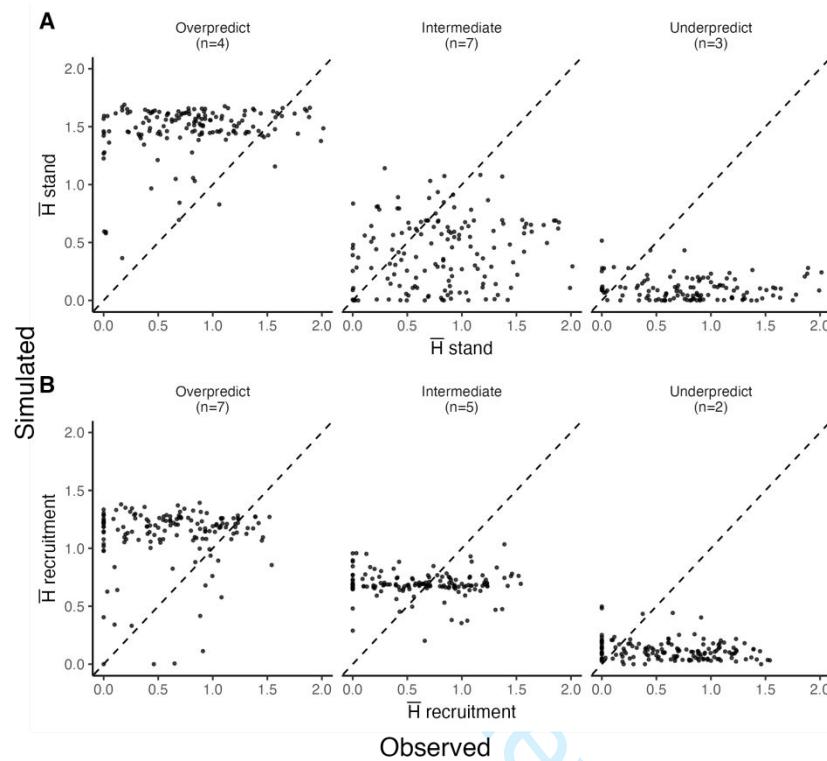
For Review Only

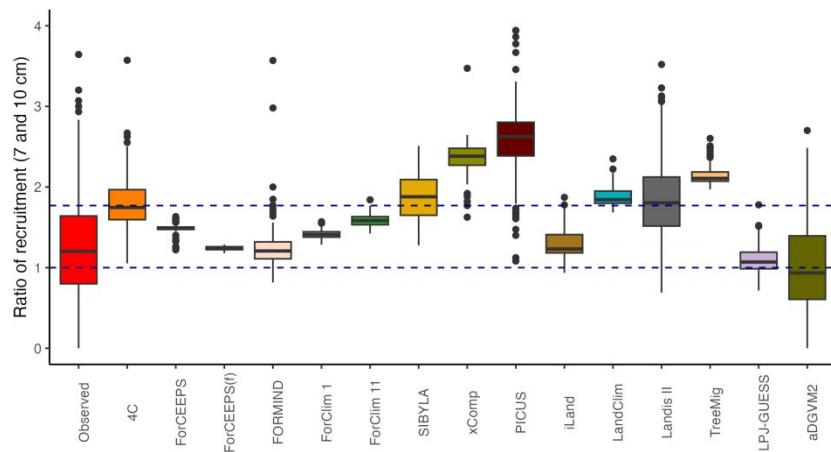
Figures

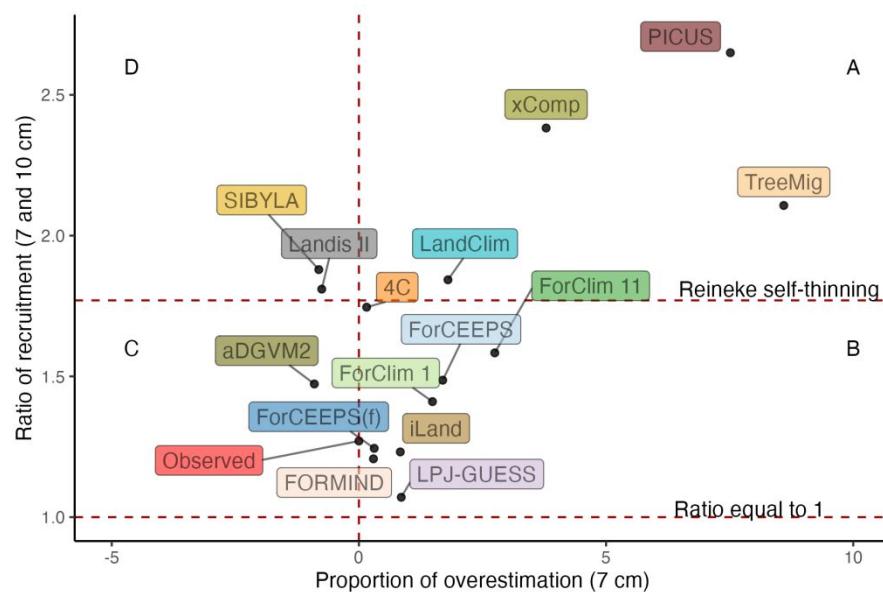


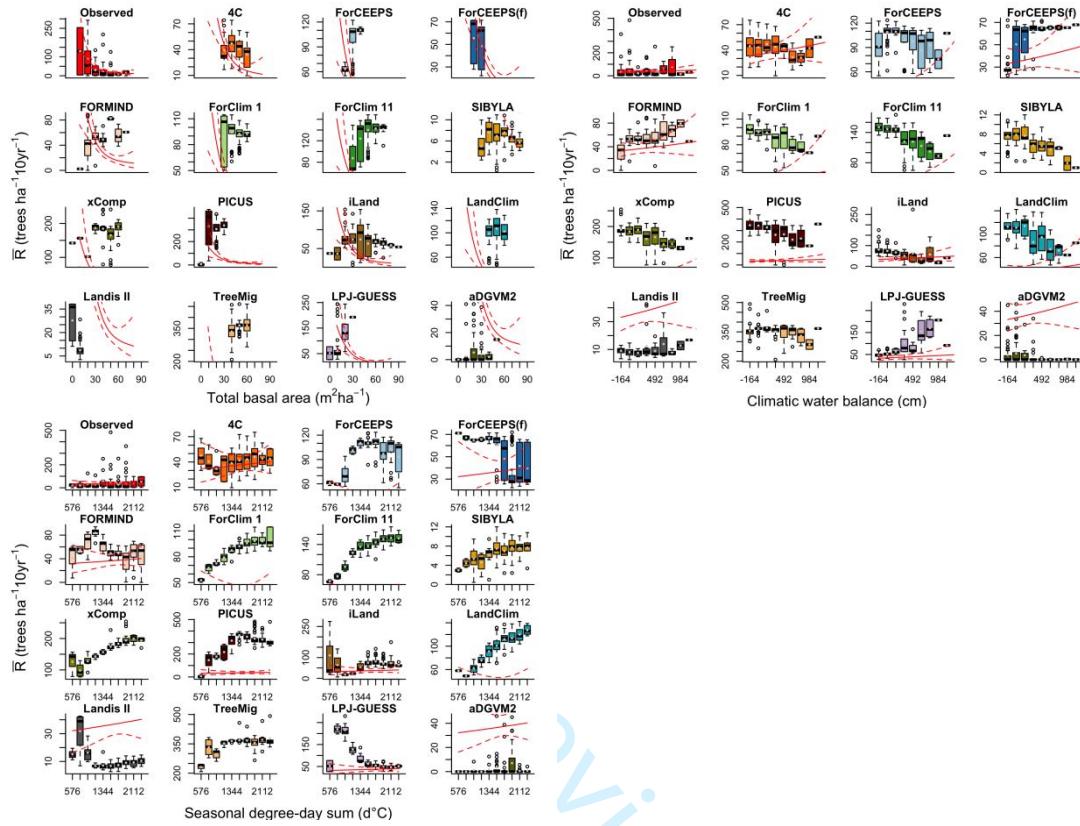


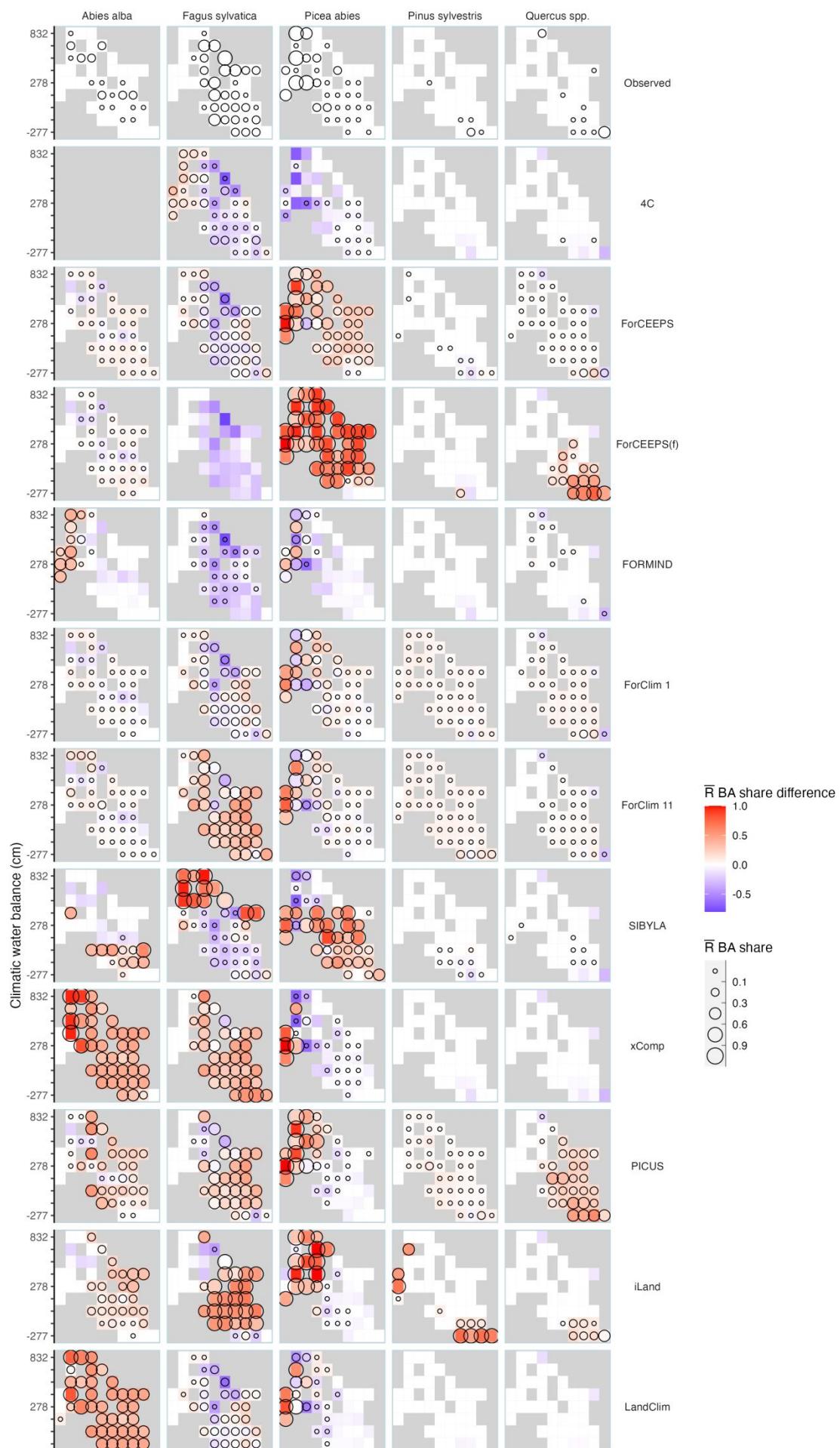












For Review Only

For Review Only