

Forecasting seed production in perennial plants: identifying challenges and charting a path forward

Valentin Journé¹, Andrew Hacket-Pain², Iris Oberklammer³, Mario B. Pesendorfer³, Michał Bogdziewicz¹

¹Forest Biology Center, Faculty of Biology, Adam Mickiewicz University, Uniwersytetu Poznańskiego 6, 61-614 Poznan, Poland.

²Department of Geography and Planning, School of Environmental Sciences, University of Liverpool, Liverpool, United Kingdom.

³Institute of Forest Ecology, Department of Forest and Soil Sciences, University of Natural Resources and Life Sciences, Vienna, Peter-Jordan-Strasse 82, A-1190 Austria

keywords: masting | forecasting | resource | climate | seed production

Corresponding authors: michalbogdziewicz@gmail.com & journe.valentin@gmail.com

ORCID:

Valentin Journé: <https://orcid.org/0000-0001-7324-7002>

Andrew Hacket-Pain: <https://orcid.org/0000-0003-3676-1568>

Iris Oberklammer: <https://orcid.org/0000-0002-6518-271X>

Mario B. Persendorfer: <https://orcid.org/0000-0002-7994-7090>

Michał Bogdziewicz: <https://orcid.org/0000-0002-6777-9034>

Twitter:

Valentin Journé: [@Valentin_Journe](https://twitter.com/Valentin_Journe); Iris Oberklammer: [@I_Uppercanyon](https://twitter.com/I_Uppercanyon) ; Mario B. Persendorfer: [@MBPesendorfer](https://twitter.com/MBPesendorfer); Michał Bogdziewicz: [@mic_bog](https://twitter.com/mic_bog)

Abstract

Interannual variability of seed production, known as masting, has far-reaching ecological impacts including effects on forest regeneration and the population dynamics of seed consumers. Because the relative timing of management and conservation efforts in ecosystems dominated by masting species often determines their success, there is a need to study masting mechanisms and develop forecasting tools for seed production. Here, we aim to establish seed production forecasting as a new branch of the discipline. We evaluate the predictive capabilities of three models - *foreMast*, ΔT , and a sequential model - designed to predict seed production in trees using a pan-European dataset of *Fagus sylvatica* seed production. The models are moderately successful in recreating seed production dynamics. The availability of high-quality data on prior seed production improved the sequential model's predictive power, suggesting that effective seed production monitoring methods are crucial for creating forecasting tools. In terms of extreme events, the models are better at predicting crop failures than bumper crops, likely because the factors preventing seed production are better understood than the processes leading to large reproductive events. We summarize the current challenges and provide a roadmap to help advance the discipline and encourage the further development of mast forecasting.

Introduction

Mast seeding or masting, highly variable and synchronized annual reproduction, is common in perennial plants (Kelly & Sork, 2002). Central to plant population dynamics, masting provides foundational but unstable resources for seed consumers, creating cycles of feast and famine in food webs (Schmidt & Ostfeld, 2008; Clark *et al.*, 2019). The pulsed nature of masting influences the success of management and conservation actions, thus making short-term predictions of seed production a powerful potential tool for practitioners (Pearse *et al.*, 2021). For instance, planning of forest regeneration actions, such as soil scarification or prescribed fires, must be timed with seed crops (Girona *et al.*, 2018), and the timing of masting in relation to disturbance may influence forest recovery, shaping management responses (Rammig *et al.*, 2007). Masting may also lead to multiyear seed scarcity that may challenge the timeframe of funded restoration projects (Woolsey *et al.*, 2018). In some species, bumper crops of seeds are extremely rare (e.g. once in several years), and seeds cannot be stored, resulting in a mismatch of planned activities and seed availability (Kettle *et al.*, 2010, 2011). Anticipatory forecasting of seed production with several-months notice, "mast cast", would therefore prepare forest agencies, managers, and other stakeholders to utilize the rare and short windows of opportunities for restoration (Kettle *et al.*, 2010, 2011). Similarly, to manage animal populations that rely on masting plants, the timing of reintroduction efforts and other interventions needs to be timed in relation to annual variation in plant reproduction (Fidler *et al.*, 2008). This includes the removal of pest species that are part of the trophic chain anchored by seeds, so that the efficacy of eradication efforts can be maximized (Köhnke *et al.*, 2020). Finally, masting affects public health near forests, as it can determine the annual pollen load for tree species, and, through indirect effects on trophic chains, the transmission of zoonotic diseases to humans (Tseng *et al.*, 2020; Bregnard *et al.*, 2021). The development of forecasts of plant reproduction and subsequent trophic cascades would thus give conservationists, land managers, and public health officials advance notice of upcoming bumper crops or crop failures. While ecological forecasting is slowly starting to enter the mainstream of the discipline (Lewis *et al.*, 2022), surprisingly sparse information exists on the current state of the art in forecasting plant reproduction (Holland & James, 2015a; Chiavetta & Marzini, 2021; Pesendorfer *et al.*, 2021; Pearse *et al.*, 2021).

74 Accurate modeling of reproduction in perennial plants has long been challenging due to the
75 large interannual and inter-individual variability inherent to masting (Vacchiano *et al.*, 2018;
76 Clark *et al.*, 2021; Journé *et al.*, 2022). This problem was further compounded by limited
77 data availability (Koenig, 2021; Clark *et al.*, 2021). However, with the recent advancement of
78 global coordination in monitoring and data synthesis, we now have access to broad taxonomic
79 data sets for plant reproduction time series (Pearse *et al.*, 2020; Clark *et al.*, 2021; Hacket-Pain
80 *et al.*, 2022). With this opportunity, we aim to develop mast forecasting ("mast cast") as a new
81 branch of the discipline. Masting may offer high potential for ecological forecasting, as the data
82 consists of extended time series with well-established abiotic drivers, potentially providing high
83 intrinsic predictability (Pennekamp *et al.*, 2019; Pearse *et al.*, 2016, 2021). Additionally, the
84 efforts to develop mast forecasting will lead to wider benefits for the discipline, as it will require
85 the formulation of quantitative predictions and challenge researchers to compare the explanatory
86 power of empirical and process-based models across different geographic regions and ecological
87 contexts (Lewis *et al.*, 2022). To take the first steps in mast forecasting, we use European beech
88 (*Fagus sylvatica*) masting across Europe as a case study to discuss the current state of the art
89 and outline a roadmap for future development. European beech is a suitable model species due
90 to its broad distribution, forest-forming nature, well-studied seed production mechanisms, and
91 exceptional data availability (Hacket-Pain *et al.*, 2022). Although our focus is on this particular
92 species, the sources of uncertainty and ways to improve masting prediction discussed in this
93 study are generalizable. Our goal is to motivate new efforts that will enhance our understanding
94 of seed production ecology and provide the tools that land managers urgently need.

95 What determines masting and how predictable is it? Currently, we understand masting as
96 a result of the interplay between weather and the internal resource dynamics of plants, and
97 their combined impact on flower and seed development (Pearse *et al.*, 2016; Pesendorfer *et al.*,
98 2021; Bogdziewicz *et al.*, 2020a; Roncé *et al.*, 2023). Weather, as an external factor, has been
99 demonstrated to impact various stages of the reproductive process, including flower initiation,
100 pollination, and seed abortion (Kon & Noda, 2007; Koenig *et al.*, 2015; Nussbaumer *et al.*, 2020;
101 Satake & Kelly, 2021). On the other hand, the endogenous effects of resource depletion (e.g.
102 carbon or nitrogen) limit seed production after large seed crops (Crone *et al.*, 2009; Sala *et al.*,
103 2012; Han & Kabeya, 2017). Thus, predictions of masting mainly rely on our understanding
104 of the links between weather and seed production, the sensitivity of masting to weather and
105 resource depletion, and the uncertainty of these drivers.

106 To understand the current state of seed production models for *F. sylvatica*, we compare two
107 published tools for predicting masting based on empirical relationships between seed production
108 and weather, the ΔT model and *foreMast*, with a mechanism-based model (hereafter "sequential
109 model") that includes additive weather effects along a sequence of phenological steps involved
110 in seed production. The ΔT model predicts seed production based on the difference in summer
111 weather between the two years leading up to seed fall (Kelly *et al.*, 2013). It does not include
112 the impact of resource availability, instead relying on the idea of an epigenetic weather memory
113 that results in elevated flower production if hot summers follow cold ones (large ΔT values)
114 (Samarth *et al.*, 2020, 2021). The ΔT forecast is utilized by New Zealand's National Predator
115 Control Program, for example in forecasting Southern beech masting at sites where native species
116 are vulnerable to the effects of mast-induced increases in the abundance of invasive predators
117 (Holland *et al.*, 2018), but it can also be applied to *F. sylvatica* (Vacchiano *et al.*, 2017). The
118 second tool, *foreMast* (Chiavetta & Marzini, 2021), is an R package developed specifically for
119 *F. sylvatica*, using data from a global time series database (Ascoli *et al.*, 2017) to forecast mast
120 events based on summer temperatures of the two years leading up to seed fall. Previous work
121 has shown that when a cold summer is followed by a hot one, beech is likely to produce a

122 bumper crop (Vacchiano *et al.*, 2017). Additionally, *foreMast* indirectly accounts for the effects
 123 of resource depletion on seed production through a parameter that decreases the probability of
 124 consecutive mast years.

125 In addition to the two existing models, we constructed the *sequential model*, which takes
 126 into account the previous year’s seed production as well as a series of weather drivers of seed
 127 production in *F. sylvatica*. Unlike ΔT and *foreMast*, the sequential model explicitly includes the
 128 direct relationship between sequential years of seed production, a crucial factor in mast seeding
 129 (Knops *et al.*, 2007; Crone & Rapp, 2014; Miyazaki *et al.*, 2014). Additionally, *F. sylvatica* seed
 130 production is not only correlated with temperatures in the past two summers (Vacchiano *et al.*,
 131 2017), but also with the weather during pollination in spring (Mund *et al.*, 2020), late spring
 132 frosts (Allevato *et al.*, 2019; Journé *et al.*, 2021), and summer drought (Nussbaumer *et al.*, 2020)
 133 (see Table 1 for a model summary). By including predictors from various time periods before
 134 the seed fall event, we hope to gain an initial understanding of the potential forecast horizon for
 135 *F. sylvatica* seed production. This approach should provide insights into which data to collect,
 136 how future models should be structured, and what experiments should be prioritized to enhance
 137 the forecasting ability of masting in this species and beyond (Dietze, 2017).

Table 1: Overview of the three models used to hindcast seed production in *Fagus sylvatica*. †
 T_0 is the seed fall year, T_1 is the year preceding the seed fall, T_2 two years preceding the seed
 fall.

Model	Coefficient type	List of predictors used
<i>foreMast</i>	Inherited from <i>foreMast</i> package	Summer precipitation (T_1, T_2), summer temperature (T_1, T_2), auto-correlation term that helps to avoid two consecutive high-seed production years.
ΔT	Site specific estimation	Summer temperature difference T_1-T_2 .
Sequential	Site specific estimation	Summer temperature (T_1, T_2), seed production T_1 , late spring frost (T_0), spring temperature (T_0) and summer moisture deficit (T_0).

138 Where are we?

139 How good are current models at predicting seed production? Which parameters have the
 140 strongest impact on predictions? How does data quality matter? To start addressing these
 141 questions, we compared the *foreMast*, ΔT , and sequential models by performing a hindcast
 142 analysis to assess their ability to recreate past seed production (Petchey *et al.*, 2015; Holland
 143 & James, 2015a). In addition to traditional metrics of model fit such as R^2 , the proportion
 144 of the explained variance, and *RMSE*, the root-mean-square error, we evaluated how well the
 145 predicted values align with observed values across the range of mast production and whether they
 146 accurately capture the interannual variability (CV - coefficient of variation) of mast production
 147 in beech populations. Finally, to investigate the effect of data quality on forecasting ability,
 148 we explored the relationship between prediction errors and site-specific covariates, such as the
 149 seed production monitoring method. With this analysis, we aim to lay the foundation for mast
 150 forecasting by exploring the potential of current models to recreate past seed production in
 151 European beech (Dietze, 2017).

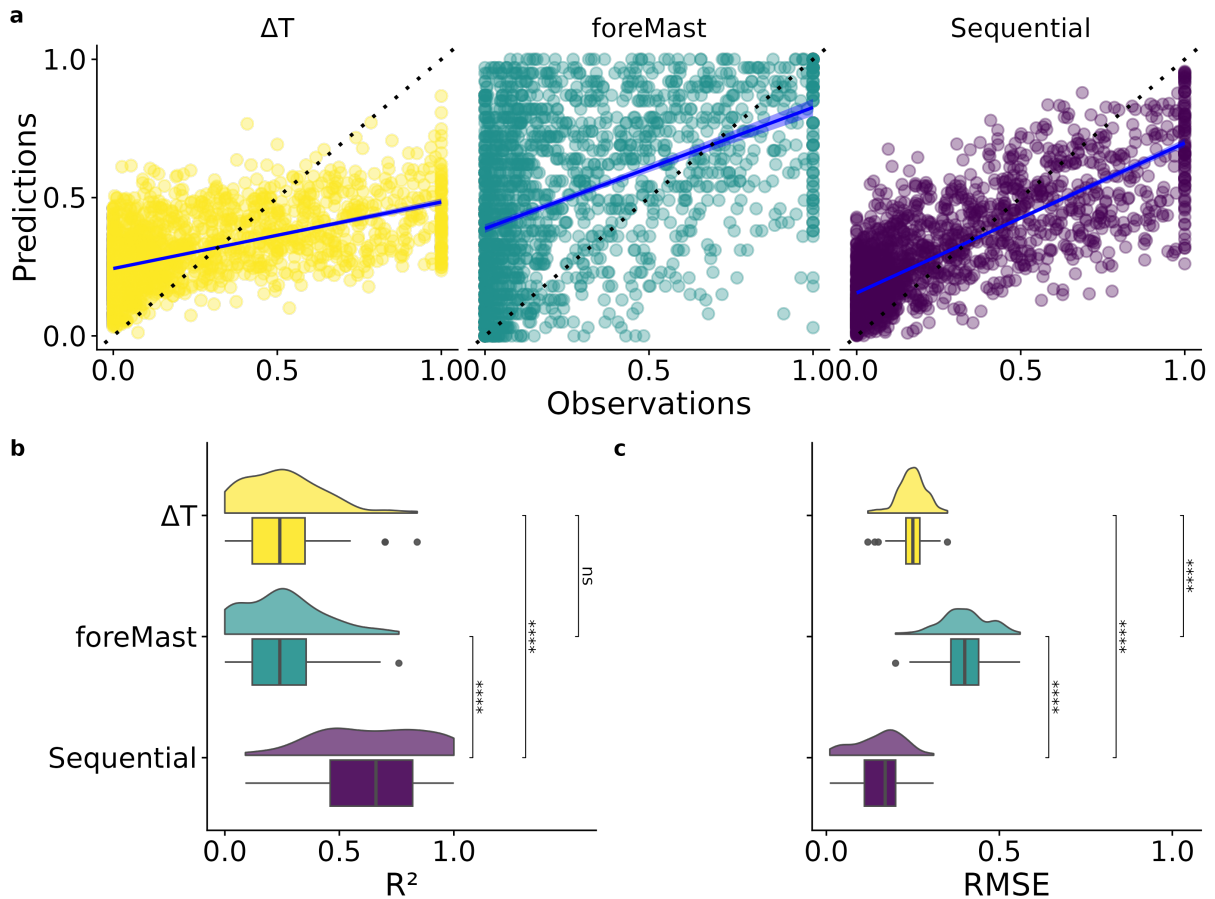


Figure 1: Goodness-of fit metrics for ΔT , *foreMast*, and the sequential model. a) Predicted vs. observed values in the three models for each site-year observation and predictions. Black dotted lines indicate a 1:1 relationship, while blue lines are regressions between predicted and observed values. b) Models R^2 , and c) Root-mean-square error (*RMSE*). We computed R^2 , i.e. the proportion of variance in the dependent variable that can be explained by the independent variables, and *RMSE*, i.e. the average difference between values predicted by a model and the actual values, for each site and combined them for each model to determine boxplot size. The asterisks show the statistical difference between each group tested with a t-test and adjusted P (**** for $P < 0.0001$, *** for $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, and n.s. for $P > 0.05$). Further metrics comparisons between the three models are presented in Figure S2 and S3.

152 To assess our ability to predict seed production in beech populations, we used 94 time series
 153 of seed production data with at least 8 years of observations from the MASTREE+ database (see
 154 Supporting Information Methods S1) (Hackett-Pain *et al.*, 2022). These data were combined
 155 with weather data for each location, starting in 1980. The *foreMast* algorithm provides masting
 156 forecasts for any given set of coordinates, while the coefficients for the ΔT and sequential
 157 models were estimated by fitting each model to each time series, yielding site-specific values
 158 for coefficients (Fig. S1). We then compared observed and predicted values for seed production
 159 in each time series. To accommodate among-site variation in methods of seed production
 160 monitoring, we scaled seed production for each site between 0 and 1 by dividing all values by
 161 the highest recorded value in the time series.

162 Models tend to over-predict seed production in years of poor seed production, and also
 163 under-predict seed production in large seeding years (Fig. 1 a). The sequential model showed
 164 higher accuracy in recreating seed production compared to the simpler *foreMast* and ΔT models
 165 (Fig. 1). The median R^2 for the sequential model was 0.65 [95% CI = 0.47, 0.82]. The median

166 R^2 for the other two models were lower, with ΔT equal to 0.26 [0.12; 0.36] and *foreMast* =
167 0.25 [0.13; 0.35]. The *RMSE* values further indicate the sequential model performed best, and
168 indicate that the ΔT model performed better than *foreMast* (Fig. 1c). The median *RMSE* for
169 the sequential model was 0.18 [95% CI = 0.12, 0.22]. The median *RMSE* for the other two
170 models were higher, with ΔT equal to 0.27 [0.24; 0.30] and *foreMast* = 0.39 [0.35; 0.44]. The
171 lower performance of the *foreMast* model was expected, as the other models were fitted locally
172 to individual time series. Meanwhile, the higher R^2 for the sequential model is associated with
173 its higher complexity, as it includes six parameters compared to only one in the ΔT model.
174 However, our goal here is not necessarily to find the most parsimonious model, but one that
175 will provide the most useful prediction of masting given the data available and the requirements
176 of the user. The weather parameters used in the sequential model are easily accessible, so the
177 increase in accuracy may come at a low cost. Still, monitoring seed production is logistically
178 challenging and some predictors in the sequential model occur during seed maturation, which
179 could reduce the potential forecast horizon (Petchey *et al.*, 2015; Pearse *et al.*, 2021). Thus, we
180 next evaluated the impact of each predictor on the performance of the sequential model.

181 To determine the impact of each parameter on the predictive power and potential forecast
182 horizon of the sequential model, we employed two approaches. First, we dropped individual
183 variables, re-fit the model, and evaluated its performance to assess the relative importance of
184 each variable. Second, we added the parameters to the model in the phenological order of their
185 occurrence, which provides information on the time frame when meaningful predictions of seed
186 production may be possible (i.e. the potential forecast horizon). Our analysis showed that the
187 previous year's summer temperature and seed production are the most significant predictors
188 of model performance (Fig. 2a,b, Fig. S4). This implies that beech seed production can be
189 predicted with good accuracy with one year's lead time, but it requires monitoring of seed
190 production (Fig. 2b). Including information on weather conditions during flowering and seed
191 maturation (spring temperature, summer drought, frost) can further improve model predictive
192 power (Fig. 2b), but it comes at the cost of a reduced forecast horizon, which may limit its
193 usefulness for certain management activities.

194 Given limited research time and resources, it is crucial to focus on factors that bring the great-
195 est improvement in predictive power. Generally, the most impactful factors on the predictability
196 of ecological processes are those that the process is highly sensitive to and those characterized by
197 high uncertainty (Dietze, 2017). As past studies and our analysis indicate, masting is impacted
198 by weather (low input uncertainty and high sensitivity) and past seed production (high input
199 uncertainty and high sensitivity). This combination implies that reliable forecasts depend on the
200 availability of high-quality seed production monitoring. To test this, we analyzed the effect of
201 different seed crop monitoring methods which vary in their accuracy on the R^2 of the models.

202 In the MASTREE+ dataset, annual seed production is estimated at the population level
203 using various methods, including seed counts - ground-counting within a certain time frame
204 (Bogdziewicz *et al.*, 2020b) -, seed traps (Bajocco *et al.*, 2021), and visual crop assessment - a
205 population-level estimate of the year's crop, such as the proportion of reproducing individuals.
206 Visual assessments are often used over large areas by foresters, e.g. by Polish State Forests
207 (Pesendorfer *et al.*, 2020). Of these methods, seed traps are expected to provide the lowest input
208 uncertainty in terms of past seed production, which leads to higher model predictive power,
209 as demonstrated by our model (Fig. S5). The median R^2 for the sequential model trained on
210 time series based on seed traps (0.74 [95% CI = 0.45, 0.87]) was much higher than those of
211 visual crop assessment (0.51, 95% CI = 0.44, 0.60) and the seed count method (0.58, 95% CI
212 = 0.49, 0.65). It is worth noting that all "seed count" time series in our data come from trees
213 in the UK beech monitoring program (Fig. S5), which was severely affected by climate change

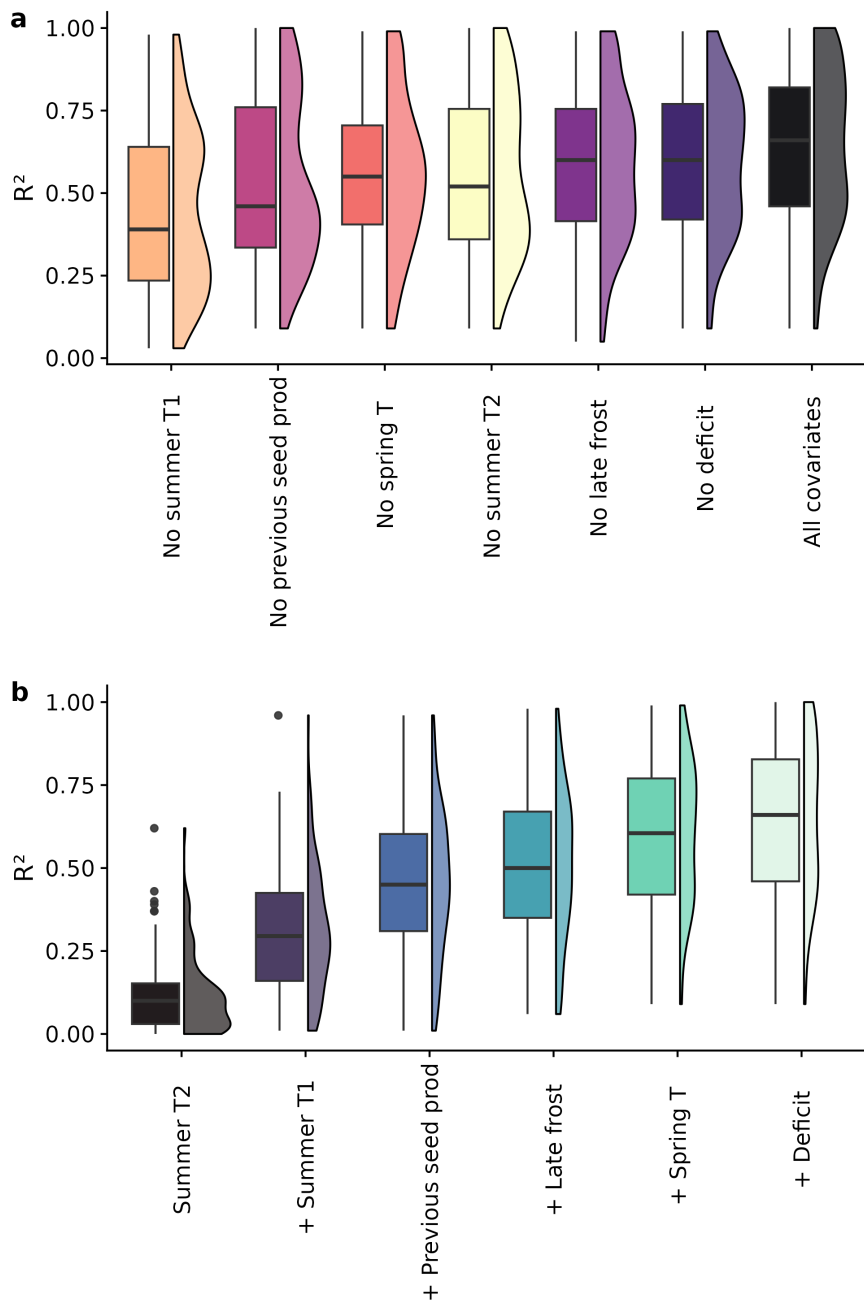


Figure 2: Effects of focal predictors on the sequential model performance (R^2). a) Each boxplot shows the sequential model R^2 with one predictor removed, b) predictors added in phenological order. We determined R^2 for each site and computed them together to determine boxplot size. An analogous figure with $RMSE$ is in Fig. S4

214 (Bogdziewicz *et al.*, 2020b). It is therefore uncertain whether the low R^2 values are due to
 215 the method's accuracy (Koenig *et al.*, 1994; Touzot *et al.*, 2018; Tattoni *et al.*, 2021) or due to
 216 changes in the climate that altered the masting processes at these sites, which merits further
 217 investigation (Bogdziewicz, 2022).

218 Forecasting large mast years and crop failures is crucial due to their extensive impact on
 219 ecosystem functioning (Pearse *et al.*, 2021). Therefore, we examined whether the models predict
 220 the data well at the extreme ends (mast and failure years). To this end, we evaluated whether the
 221 predicted values are low enough (in failure years) or high enough (in mast years). We defined

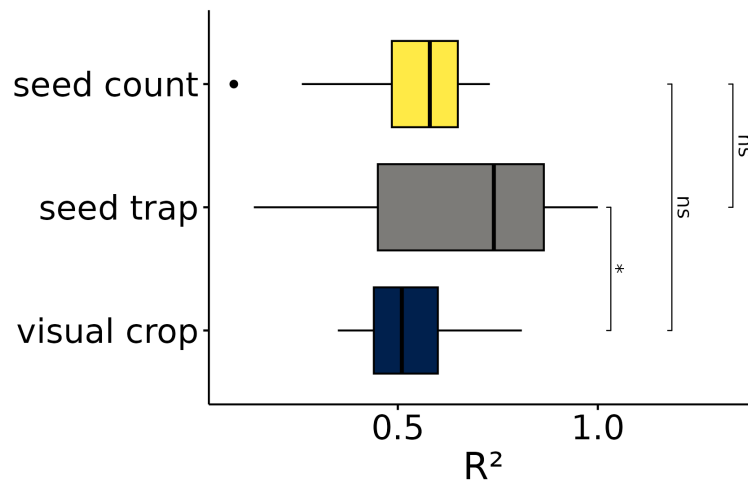


Figure 3: Effects of seed production estimation method on the sequential model R^2 . The methods are seed count (ground counting within a given time frame), monitoring with seed traps, and visual crop assessment. The asterisks represent significant differences between each group tested with a t-test (**** for $P < 0.0001$, *** for $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, and n.s. for $P > 0.05$). An analogous figure with $RMSE$ is in Fig. S6.

222 mast years as years when seed production was higher than the site-specific 80th centile ($9.64\% \pm$
 223 5.19 SD site-years), and failure years as below the 20th centile ($62.8\% \pm 17.5$ SD site-years).
 224 Importantly, that categorization was done only once the predictions of the models were obtained,
 225 not for fitting the models. We note that dividing seed production into mast and non-mast is an
 226 ecologically false dichotomy in the sense that the distribution of seed production is not binomial,
 227 and the "mast year" definition is subjective. Nonetheless, such categorization can be useful from
 228 the forecasting/management perspective and is often used in forestry (Pearse *et al.*, 2021).

229 Our evaluation revealed that the three models had limited success in accurately predicting
 230 extreme seeding events. In our dataset, large mast years occurred in 9.64% of the years. The ΔT
 231 model correctly predicted these years in only 0.19% of cases, but got false negative in 9.45%
 232 of the observed mast years. Equivalent performance for *foreMast* was 4.68% true positive and
 233 4.97% false negative, and for the sequential model it was 3.63% true positive and 6.01% false
 234 negative. On the other hand, the models performed better in predicting crop failures (defined
 235 as seed production below 0.2), although still far from satisfactory (Fig. 4a). In our dataset,
 236 crop failures occurred in 62.8% of the years. The ΔT model correctly predicted these failures
 237 in 19.9% of the years and predicted 42.9% of false negatives. Similar results were observed
 238 for the *foreMast* model, with correct predictions of 17.4% and failed predictions with 45.4% of
 239 false negative. The sequential model performed the best in this aspect, with correct predictions
 240 of 42.8% and failed predictions with false negatives in 20.0% (Table 2).

241 Interannual variation of seed production is the defining feature of masting. Thus, we further
 242 examined how the interannual variation of predicted values recreates the observed variation.
 243 Despite the better performance in predicting low seed production years, the overall poor ability
 244 of the models to predict large seed crops leads to low estimates of the variability of masting time
 245 series (Fig. 4b). A similar finding was reported in a study of seed production in valley oaks
 246 (*Quercus lobata*), where historical crop failures were accurately modeled, but the extent of seed
 247 production in mast years was not well captured (Pesendorfer *et al.*, 2016).

248 Assessing the ability of the models to predict extreme seeding events (large mast and failure
 249 years) in addition to their performance based on R^2 provides valuable insights into models'
 250 effectiveness. Both mast and failure years are underpredicted, with especially poor performance

Table 2: Prediction of mast and failure years by the three evaluated models. For each time series, we scaled seed production to a range of 0 - 1, and defined mast years as years when seed production was above 0.8, whereas failures as below 0.2. Based on this arbitrary threshold, mast years occurred in $\sim 9.64\%$, and failure years occurred in $\sim 62.8\%$ of years. We reported average values with standard deviation (SD) in percent.

	Mast years	Failure years
<i>True positive (hit)</i>		
ΔT	0.19% \pm 1.33	19.9% \pm 17.9
<i>foreMast</i>	4.68% \pm 4.42	17.4% \pm 5.55
Sequential	3.63% \pm 5.78	42.8% \pm 21.6
<i>False positive (type one error)</i>		
ΔT	0% \pm 0	1.18% \pm 3.05
<i>foreMast</i>	14.0% \pm 6.1	1.92% \pm 3.97
Sequential	0.42% \pm 1.52	1.94% \pm 3.71
<i>False negative (type two error)</i>		
ΔT	9.45% \pm 5.17	42.9% \pm 16.6
<i>foreMast</i>	4.97% \pm 4.74	45.4% \pm 15.7
Sequential	6.01% \pm 5.10	20.0% \pm 15.0

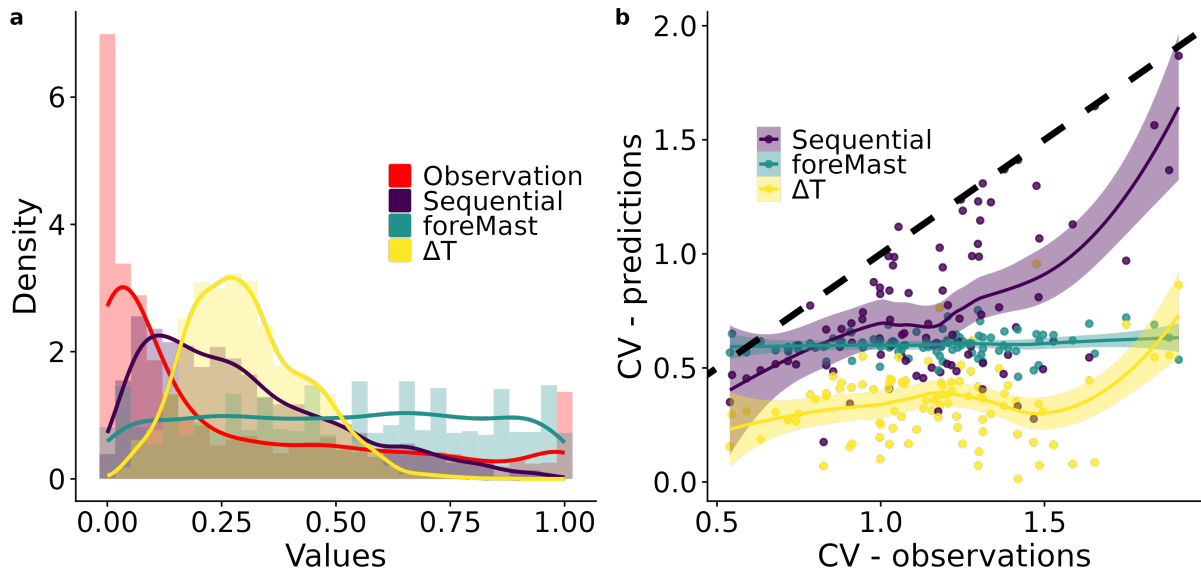


Figure 4: Distribution and variation of observed and predicted seed production. a) Histograms with density curves of all observed and predicted annual seed production. b) Coefficient of variation (CV) calculated based on observed and predicted values of annual seed production at each site. We computed CV by dividing standard deviation (SD) by the mean value of seed production of the time series for each site. Predictions are based on the three evaluated models, i.e. ΔT , *foreMast*, and sequential model. We computed CV based on annual seed predictions, and from scaled observations, for each sites (Sequential, $R^2 = 0.29$; *foreMast*, $R^2 = 0.02$; ΔT , $R^2 = 0.05$). For the three models in b) we fitted a loess regression to visualize trends.

251 in the prediction of mast years. The sequential model predicts failure years better, partially due to
 252 its inclusion of the previous year's seed production. That model accuracy in predicting failures
 253 decreases when the previous year's seed production is removed (correctly predicted true positive
 254 failure years: 37.2%). The sequential model also performs better in predicting failure years
 255 when trained solely on seed trap data, which is associated with lower uncertainty (correctly

256 predicted true positive failure years: 47.2%). These findings support the idea that resource
257 depletion plays an important role in seed production, making seed failure after high-seed years
258 predictable (Crone *et al.*, 2009; Abe *et al.*, 2016), while the mechanisms driving high-seed years
259 are less clear or more stochastic (Schermer *et al.*, 2020), highlighting this as a priority for future
260 research (Bogdziewicz, 2022).

Box 1 :**Hindcasting vs forecasting.**

Analysis in our study is based on hindcast analysis that assessed models' ability to recreate past seed production. Actual forecasting will require models to be trained on existing data and predict unobserved future seed production, which will likely result in reduced performance. To illustrate that, we used a cross-validation procedure with the sequential model, based on a random data-split (Yates *et al.*, 2023). For each site, we trained the model with 70% of randomly selected observations, and repeated this process five times. The R^2 was evaluated with the remaining 30% of observations across the 10-fold block selection for a model trained on 70% of the observations. With cross-validation, observations can be used for training or validation only, whereas a model trained with the full dataset would include these 30% of observations for both model training and validation. Training data on a subset reduces the sequential model accuracy in predicting both failure and mast years (Fig. 5). On average, the models trained with split data achieved a lower mean R^2 at 0.41 [CI 97.5% = 0.12, 0.65] compared to a model trained with all observations ($R^2 = 0.65$ [0.42, 0.94], Fig. 5b). The take-home message is that forecasting masting shows great early promise, but substantial efforts will be required before we can predict masting with satisfying accuracy.

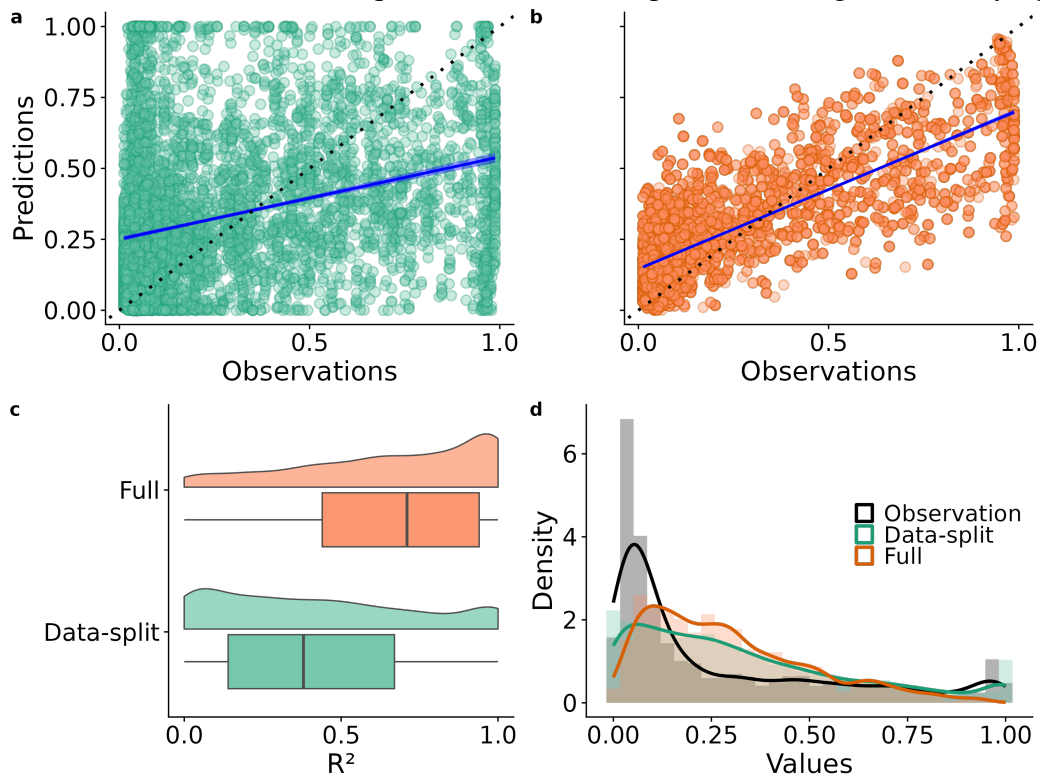


Figure 5: Comparison between cross-validation (data-split) and full data for model training. Predicted vs observed values for a model with (a) data-split and (b) a model trained on full data. Black dotted lines indicate a 1:1 relationship, while blue lines are regressions between predicted and observed values, each dot is a site-year. c) The R^2 from the model trained on 70% of the dataset or trained on the full dataset; an analogous figure with $RMSE$ is given in Fig S7. d) Histogram and density of observed values, values predicted in cross-validation (70% of the data used for calibration, 30% for the validation), and values predicted by the model trained on the full data. Note that we used 30% of the observations for each site randomly sampled ten times, which resulted in a different sample compared to Fig. 1.

261 **What is next?**

262 Our results, based on one of the best-studied systems in the field, suggest that much work is
263 needed to develop reliable tools for forecasting seed production in masting plants. The abilities
264 of models to recreate past seeding were limited, even in predicting the data on which they were
265 trained (Box 1). However, the path forward has become clearer. The sequential model, which
266 considers a broader range of factors affecting seed production compared to the other two models,
267 shows promising accuracy. Despite this, the models' inability to accurately predict bumper crops
268 remains a crucial limitation. One potential solution could be to transition the sequential model
269 from an additive to an interactive configuration (Kelly *et al.*, 2013; Vacchiano *et al.*, 2017). On
270 the other hand, the ability to predict crop failures is much improved compared to other models,
271 likely due to the deterministic impact of seed production on future reproductive investment.

272 Forecasting in ecology depends on our ability to identify key parameters that drive the system
273 and how accurately we can measure these parameters (uncertainty) (Dietze, 2017). To achieve
274 accurate forecasts, it is crucial to identify the right parameters and fine-tune them effectively. For
275 instance, enhancing our understanding of how weather variability influences seed production
276 can lead to substantial improvement in the quality of masting forecasts. Currently, correlations
277 between masting and weather are often based on fixed time frames, such as mean maximum
278 temperatures in June-July (Vacchiano *et al.*, 2017). However, the timing of these effects can
279 vary depending on local climates and current year conditions, similar to other phenological
280 events like bud burst or flowering (Fu *et al.*, 2015; Zohner *et al.*, 2018; Nussbaumer *et al.*, 2018).
281 Using a site-specific moving window analysis can help pinpoint the right weather correlate and
282 its timing (Bogdziewicz *et al.*, 2023). While in European beech the cues change little among
283 populations (Vacchiano *et al.*, 2017; Bogdziewicz *et al.*, 2023), in other species the weather
284 correlates and underlying mechanisms can vary much more among populations (Bogdziewicz
285 *et al.*, 2020c; Fleuret *et al.*, 2023). The challenge ahead lies in developing general rules that can
286 connect local variation to a global model.

287 A related problem are the unknown mechanisms that connect masting and weather variation,
288 as these links are often only assumed after the correlations are established (Crone & Rapp,
289 2014; Bogdziewicz *et al.*, 2020a). The phenological synchrony hypothesis stands out compared
290 to other weather cues with a developed theory and experiments on how weather affects seed
291 crops through pollination efficiency (Koenig *et al.*, 2015). However, even in this case, the
292 exact mechanism behind flowering synchrony, a key factor in pollination efficiency, remains
293 unclear. Different theories propose different temperature sensitivities, with the photoperiod-
294 sensitivity hypothesis suggesting temperature affects flowering months before it occurs (Chuine,
295 2010; Bogdziewicz *et al.*, 2020c), while the micro-climatic hypothesis posits that temperature
296 during flowering is crucial (Koenig *et al.*, 2015). The effects of warm spring weather on seed
297 production can be also explained by higher pollen dispersal in such conditions (Schermer *et al.*,
298 2019). This disagreement on the relevant period makes it difficult to identify the timing of
299 maximum sensitivity, adding complexity to the task of constructing reliable masting forecasting
300 tools.

301 Improving the accuracy of forecasts can also be achieved by reducing uncertainty in param-
302 eters. The uncertainty of weather parameters is generally low as past measurements are mainly
303 used. However, for parameters with high sensitivity, which increases the cost of uncertainty,
304 weather can be measured directly at the focal site instead of relying on data from nearby me-
305 teorological stations. To assess the benefits of local measurements, case studies are needed.
306 Another important source of uncertainty is the impact of past seed production on current-year
307 reproductive investment. Therefore, determining the most effective method of seed monitoring

308 at relevant spatial scales is a priority. The method chosen can vary based on factors such as plant
309 species (e.g. the spatial scale of seed production synchrony, the importance of post-seeding
310 resource depletion on next year's seed production) (Touzot *et al.*, 2018; Tattoni *et al.*, 2021), and
311 the objective of the forecast. Further research into the physiology of resource depletion following
312 seed production holds promise for improvement, including examining how stand age/mean size
313 or local factors (soil type, competition) affect the limiting role of post-mast resource depletion
314 on subsequent year's seed production (Pesendorfer *et al.*, 2020; Wion *et al.*, 2023), or whether
315 extreme mast years can result in two consecutive mast failures.

316 **Conclusions**

317 The vast array of masting strategies across species, and potentially even within species (Fleurot
318 *et al.*, 2023), may make it difficult to develop a universal model of masting. To address this
319 challenge, stakeholders, including experts on masting and those who use such information,
320 should collaborate to design research projects that focus on specific goals, rather than relying
321 solely on lessons learned from other systems (Dietze, 2017). However, a general understanding
322 of masting is crucial in providing a framework to work toward specific applications. To facilitate
323 this, we summarize the key points in Box 2 .

324 It is worth mentioning that although the sequential model outperformed in our analysis, this
325 does not necessarily render other models ineffective. For instance, our analysis was focused on
326 European beech, but the ΔT model has already proven successful in forecasting seed production
327 in Southern beech (Holland & James, 2015a). Our intention is not to suggest that certain models
328 are superior or inferior, but rather to emphasize that existing models show potential, require
329 enhancements in key areas, and their effectiveness will depend on specific applications.

330 A crucial priority is to determine the intrinsic predictability of seed production time series.
331 In other words, how predictable is masting, and what is the achievable level of accuracy for
332 the models? It is important to understand if the system is inherently predictable to evaluate the
333 quality of the chosen model (Pennekamp *et al.*, 2019). The impact of chaotic dynamics, the effect
334 of changing initial conditions on model results, should also be considered when assessing the
335 models (Rogers *et al.*, 2022). Improving our understanding of what can be predicted and how to
336 improve those predictions is essential for making ecology more relevant to policy, management,
337 and decision-making (Clark *et al.*, 2001; Dietze, 2017).

Box 2 :

The Next Steps: A Roadmap for Masting Forecasting. To advance masting forecasting the following research areas are identified as essential or promising. This roadmap aims to provide guidance for future efforts in this field and to help advance our understanding of masting dynamics and its predictability.

- **Predicting mast years and failures.** Classifying seed production into mast and non-mast years is subjective and does not accurately represent the complex ecological processes involved. However, extreme events play a significant role in management decisions. The limitations of current models in accurately forecasting these stochastic large events highlight the need for further research into the mechanisms that drive mast years.
- **Physiology of seed production.** Promising results have been achieved using sequential models, which are based on established correlative mechanisms. This suggests that a deeper understanding of the mechanisms behind seed production and weather correlations, as well as the factors that limit post-mast seed production (such as the resource allocation to flowering and identification of resources that actually limit reproduction), could further improve the effectiveness of these models.
- **Model types.** The development of models for masting forecasting should follow an iterative process, in which model performance is constantly evaluated, refined, and re-evaluated (Dietze *et al.*, 2018). The selection and validation of models are crucial steps in this process and can be improved by identifying and separating sources of uncertainty in the forecast. For example, Heilman *et al.* (2022) separately evaluated the uncertainty from initial conditions, parameters, drivers, and processes. Within the iterative cycle, comparing the explanatory power and predictive ability of both mechanism-based and empirical models (Lewis *et al.*, 2022), as well as artificial intelligence approaches such as deep learning (Christin *et al.*, 2019), can be a promising approach. Our sequential model showcases the advantage of a mechanism-informed empirical approach, which adds complexity but significantly improves hindcasting accuracy. Alternative routes include other mechanism-based models, such as the resource budget model of masting (Satake & Iwasa, 2000; Holland & James, 2015b; Schermer *et al.*, 2020), or phenomenological models (Lebourgeois *et al.*, 2018).
- **Identifying and incorporating rare veto factors.** Incorporating rare and difficult-to-quantify veto events presents a challenge in masting forecasting. Such events, such as late spring frosts or extreme droughts, can have significant localized impacts by killing flowers or causing fruit abortion (i.e. high sensitivity). However, these events are rare and only have significance if the masting process has already been initiated (i.e. previous processes in the phenological sequence that leads to large mast years have happened). Measuring their effects on masting is challenging due to the limited data available and their non-linear, threshold-based nature. Despite these difficulties, the predicted increase in the frequency of such events in the future suggests they may play an increasingly important role in driving masting.
- **Intrinsic vs. extrinsic drivers of masting** Certain species show bimodal seed production patterns, which are only altered due to vetoes. For example, silver fir *Abies alba* shows strong biannual reproduction unless a late frost acts as a veto (Owens & Morris, 1998). Thus, by investigating the cyclicity and autocorrelation of reproduction across

species in different ecological contexts, we may learn more about the role of intrinsic and extrinsic drivers of reproductive variation.

- **Informed model selection by users.** Different models have varying strengths when it comes to forecasting masting, and the ideal model may depend on the specific needs and requirements of the user. For instance, the relative significance of incorrectly predicted mast events (false positives) compared to missed mast years (false negatives) can influence the choice of a model. To facilitate informed decision-making, it is crucial to establish a stakeholder-driven framework for evaluating models, which will allow users to choose the best model for their specific needs.
- **Non-stationarity and predictability of seed production.** To determine the realistic level of forecast accuracy of the models, it is important to assess the intrinsic predictability of the seed production time series. This includes determining the extent to which the masting system is inherently predictable, and evaluating the chosen model quality in relation to the system's intrinsic predictability (Pennekamp *et al.*, 2019). Comparing the intrinsic predictability across different masting species would provide valuable insights and be a fascinating research endeavor.
- **Climate change.** The changing climate is having a significant impact on seed production patterns (Bogdziewicz, 2022) and therefore on masting forecasting. For instance, the sensitivity of seed production in European beech to summer weather temperatures is declining, reducing the sensitivity of masting to a key predictor (Bogdziewicz *et al.*, 2021). It is crucial to understand the consequences of these changes in climate on other parameters and how they translate into the predictability of masting. This will help to improve the forecasting of production patterns in the face of ongoing climate change.
- **Data quality, parameters uncertainty, seed production monitoring.** Determining the measurement of seed production is crucial for future masting forecasting efforts. The question arises as to what spatial scale seed production should be measured and how many individuals should be included in the measurement. With the presence of spatial synchrony in seed production (LaMontagne *et al.*, 2020; Wion *et al.*, 2020), forecasts may not necessarily rely on direct observations from the targeted population. It is imperative to conduct studies that will establish a cost-effective seed production monitoring program, one that will strike a balance between the logistical challenges of monitoring and the management objectives.
- **Temporal vs spatial prediction.** The analysis presented here is largely focused on temporal predictions, i.e. forecasting future seed production. Another potentially relevant approach is spatial prediction, i.e. predicting seed production at sites where observations are missing. The efficacy of such prediction will hinge on the degree of spatial synchrony in seed production exhibited by the species under scrutiny (Bogdziewicz *et al.*, 2023), but perhaps also on stand characteristics such as soils or topography.
- **Rapid data access.** To develop accurate forecasts, timely and accessible seed production data is essential. Platforms such as the Ecological Forecasting Initiative (<https://ecoforecast.org/>), which focuses on fine-tuning ecological forecasting through rapid data sharing, can serve as a model for seed production research. The establishment of similar platforms will greatly benefit the development and evaluation of masting forecasting.

References

- 338
- 339 Abe, T., Tachiki, Y., Kon, H., Nagasaka, A., Onodera, K., Minamino, K. *et al.* (2016). Pa-
340 rameterisation and validation of a resource budget model for masting using spatiotemporal
341 flowering data of individual trees. *Ecology Letters*, 19, 1129–1139.
- 342 Allevato, E., Saulino, L., Cesarano, G., Chirico, G.B., D’Urso, G., Bolognesi, S.F. *et al.* (2019).
343 Canopy damage by spring frost in european beech along the apennines: effect of latitude,
344 altitude and aspect. *Remote Sensing of Environment*, 225, 431–440.
- 345 Ascoli, D., Maringer, J., Hacket-Pain, A., Conedera, M., Drobyshev, I., Motta, R. *et al.* (2017).
346 Two centuries of masting data for European beech and Norway spruce across the European
347 continent. *Ecology*, 98, 1473.
- 348 Bajocco, S., Ferrara, C., Bascietto, M., Alivernini, A., Chirichella, R., Cutini, A. *et al.* (2021).
349 Characterizing the climatic niche of mast seeding in beech: Evidences of trade-offs between
350 vegetation growth and seed production. *Ecological Indicators*, 121, 107139.
- 351 Bogdziewicz, M. (2022). How will global change affect plant reproduction? a framework for
352 mast seeding trends. *New Phytologist*, 234, 14–20.
- 353 Bogdziewicz, M., Ascoli, D., Hacket-Pain, A., Koenig, W.D., Pearse, I., Pesendorfer, M. *et al.*
354 (2020a). From theory to experiments for testing the proximate mechanisms of mast seeding:
355 an agenda for an experimental ecology. *Ecology Letters*, 23, 210–220.
- 356 Bogdziewicz, M., Hacket-Pain, A., Kelly, D., Thomas, P.A., Lageard, J. & Tanentzap, A.J.
357 (2021). Climate warming causes mast seeding to break down by reducing sensitivity to
358 weather cues. *Global Change Biology*, 27, 1952–1961.
- 359 Bogdziewicz, M., Kelly, D., Thomas, P.A., Lageard, J.G. & Hacket-Pain, A. (2020b). Climate
360 warming disrupts mast seeding and its fitness benefits in european beech. *Nature Plants*, 6,
361 88–94.
- 362 Bogdziewicz, M., Szymkowiak, J., Bonal, R., Hacket-Pain, A., Espelta, J.M., Pesendorfer, M.
363 *et al.* (2020c). What drives phenological synchrony? warm springs advance and desynchronize
364 flowering in oaks. *Agricultural and Forest Meteorology*, 294, 108140.
- 365 Bogdziewicz, M.C., Journe, V., Hacket-Pain, A. & Szymkowiak, J. (2023). Mechanisms driving
366 interspecific variation in regional synchrony of trees reproduction. *Ecology Letters*.
- 367 Bregnard, C., Rais, O. & Voordouw, M.J. (2021). Masting by beech trees predicts the risk of
368 lyme disease. *Parasites Vectors*, 14, 168.
- 369 Chiavetta, U. & Marzini, S. (2021). foreMast: an R package for predicting beech (*Fagus sylvatica*
370 L.) masting events in European countries. *Annals of Forest Science*, 78, 1–10.
- 371 Christin, S., Éric Hervet & Lecomte, N. (2019). Applications for deep learning in ecology.
372 *Methods in Ecology and Evolution*, 10, 1632–1644.
- 373 Chuine, I. (2010). Why does phenology drive species distribution? *Philosophical Transactions*
374 *of the Royal Society B: Biological Sciences*, 365, 3149–3160.

- 375 Clark, J.S., Andrus, R., Aubry-Kientz, M., Bergeron, Y., Bogdziewicz, M., Bragg, D.C. *et al.*
376 (2021). Continent-wide tree fecundity driven by indirect climate effects. *Nature Communi-*
377 *cations*, 12, 1–11.
- 378 Clark, J.S., Carpenter, S.R., Barber, M., Collins, S., Dobson, A., Foley, J.A. *et al.* (2001).
379 Ecological forecasts: An emerging imperative. *Science*, 293, 657–660.
- 380 Clark, J.S., Nuñez, C.L. & Tomasek, B. (2019). Foodwebs based on unreliable foundations:
381 spatiotemporal masting merged with consumer movement, storage, and diet. *Ecological*
382 *Monographs*, 89, 1–24.
- 383 Crone, E.E., Miller, E. & Sala, A. (2009). How do plants know when other plants are flowering?
384 resource depletion, pollen limitation and mast-seeding in a perennial wildflower. *Ecology*
385 *Letters*, 12, 1119–1126.
- 386 Crone, E.E. & Rapp, J.M. (2014). Resource Depletion, Pollen Coupling, and the Ecology of
387 Mast Seeding. *Ann. N.Y. Acad. Sci.*, 1322, 21–34.
- 388 Dietze, M.C. (2017). Prediction in ecology: A first-principles framework. *Ecological Applica-*
389 *tions*, 27, 2048–2060.
- 390 Dietze, M.C., Fox, A., Beck-johnson, L.M., Betancourt, J.L. & Hooten, M.B. (2018). Iterative
391 near-term ecological forecasting : Needs , opportunities , and challenges. *Proceedings of the*
392 *National Academy of Sciences*, 115, 1424–1432.
- 393 Fidler, A.E., Lawrence, S.B., McNatty, K.P., Fidler, A.E., Lawrence, S.B. & McNatty, K.P.
394 (2008). Viewpoint. an hypothesis to explain the linkage between kakapo (*Strigops habroptilus*)
395 breeding and the mast fruiting of their food trees. *Wildlife Research*, 35, 1–7.
- 396 Fleurot, E., Lobry, J.R., Boulanger, V., Debias, F., Mermet-Bouvier, C., Caignard, T. *et al.*
397 (2023). Oak masting drivers vary between populations depending on their climatic environ-
398 ments. *Current Biology*.
- 399 Fu, Y.H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G. *et al.* (2015). Declining global
400 warming effects on the phenology of spring leaf unfolding. *Nature*, 526, 104–107.
- 401 Girona, M.M., Lussier, J.M., Morin, H. & Thiffault, N. (2018). Conifer regeneration after
402 experimental shelterwood and seed-tree treatments in boreal forests: finding silvicultural
403 alternatives. *Frontiers in Plant Science*, 9.
- 404 Hacket-Pain, A., Foest, J.J., Pearse, I.S., LaMontagne, J.M., Koenig, W.D., Vacchiano, G. *et al.*
405 (2022). MASTREE+: Time-series of plant reproductive effort from six continents. *Global*
406 *Change Biology*, 28, 3066–3082.
- 407 Han, Q. & Kabeya, D. (2017). Recent developments in understanding mast seeding in relation
408 to dynamics of carbon and nitrogen resources in temperate trees. *Ecological Research*, 32,
409 771–778.
- 410 Heilman, K.A., Dietze, M.C., Arizpe, A.A., Aragon, J., Gray, A., Shaw, J.D. *et al.* (2022).
411 Ecological forecasting of tree growth: Regional fusion of tree-ring and forest inventory data
412 to quantify drivers and characterize uncertainty. *Global Change Biology*, 28, 2442–2460.

- 413 Holland, E.P., Binny, R.N. & James, A. (2018). Optimal control of irrupting pest populations in
414 a climate-driven ecosystem. *PeerJ*, 6, e6146.
- 415 Holland, E.P. & James, A. (2015a). Assessing the efficacy of population-level models of mast
416 seeding. *Theoretical Ecology*, 8, 121–132.
- 417 Holland, E.P. & James, A. (2015b). Assessing the efficacy of population-level models of mast
418 seeding. *Theoretical Ecology*, 8, 121–132.
- 419 Journé, V., Andrus, R., Aravena, M.C., Ascoli, D., Berretti, R., Berveiller, D. *et al.* (2022).
420 Globally, tree fecundity exceeds productivity gradients. *Ecology Letters*, 25, 1471–1482.
- 421 Journé, V., Caignard, T., Hacket-Pain, A. & Bogdziewicz, M. (2021). Leaf phenology correlates
422 with fruit production in european beech (*Fagus sylvatica*) and in temperate oaks (*Quercus*
423 *robur* and *Quercus petraea*). *European Journal of Forest Research*, 140, 733–744.
- 424 Kelly, D., Geldenhuis, A., James, A., Penelope Holland, E., Plank, M.J., Brockie, R.E. *et al.*
425 (2013). Of mast and mean: Differential-temperature cue makes mast seeding insensitive to
426 climate change. *Ecology Letters*, 16, 90–98.
- 427 Kelly, D. & Sork, V.L. (2002). Mast Seeding in Perennial Plants: Why, How, Where? *Annual*
428 *Review of Ecology and Systematics*, 33, 427–447.
- 429 Kettle, C.J., Ghazoul, J., Ashton, P., Cannon, C.H., Chong, L., Diway, B. *et al.* (2011). Seeing
430 the fruit for the trees in borneo. *Conservation Letters*, 4, 184–191.
- 431 Kettle, C.J., Ghazoul, J., Ashton, P.S., Cannon, C.H., Chong, L., Diway, B. *et al.* (2010). Mass
432 fruiting in borneo: A missed opportunity. *Science*, 330, 584.
- 433 Knops, J.M., Koenig, W.D. & Carmen, W.J. (2007). Negative correlation does not imply a
434 tradeoff between growth and reproduction in california oaks. *Proceedings of the National*
435 *Academy of Sciences of the United States of America*, 104, 16982–16985.
- 436 Koenig, W.D. (2021). A brief history of masting research. *Philosophical Transactions of the*
437 *Royal Society B: Biological Sciences*, 376.
- 438 Koenig, W.D., Knops, J.M., Carmen, W.J. & Pearse, I.S. (2015). What drives masting? the
439 phenological synchrony hypothesis. *Ecology*, 96, 184–192.
- 440 Koenig, W.D., Knops, J.M., Carmen, W.J., Stanback, M.T. & Mumme, R.L. (1994). Estimating
441 acorn crops using visual surveys. *Canadian Journal of Forest Research*, 24, 2105–2112.
- 442 Kon, H. & Noda, T. (2007). Experimental investigation on weather cues for mast seeding of
443 *fagus crenata*. *Ecological Research*, 22, 802–806.
- 444 Köhnke, M.C., Binny, R.N., Holland, E.P. & James, A. (2020). The necessity of tailored control
445 of irrupting pest populations driven by pulsed resources. *Theoretical Ecology*, 13, 261–275.
- 446 LaMontagne, J.M., Pearse, I.S., Greene, D.F. & Koenig, W.D. (2020). Mast seeding patterns
447 are asynchronous at a continental scale. *Nature Plants*, 6, 460–465.
- 448 Lebourgeois, F., Delpierre, N., Dufrêne, E., Cecchini, S., Macé, S., Croisé, L. *et al.* (2018).
449 Assessing the roles of temperature, carbon inputs and airborne pollen as drivers of fructifi-
450 cation in european temperate deciduous forests. *European Journal of Forest Research*, 137,
451 349–365.

- 452 Lewis, A.S.L., Rollinson, C.R., Allyn, A.J., Ashander, J., Brodie, S., Brookson, C.B. *et al.*
453 (2022). The power of forecasts to advance ecological theory. *Methods in Ecology and*
454 *Evolution*, 00, 1–11.
- 455 Miyazaki, Y., Maruyama, Y., Chiba, Y., Kobayashi, M.J., Joseph, B., Shimizu, K.K. *et al.* (2014).
456 Nitrogen as a key regulator of flowering in *Fagus crenata*: understanding the physiological
457 mechanism of masting by gene expression analysis. *Ecology Letters*, 17, 1299–1309.
- 458 Mund, M., Herbst, M., Knohl, A., Matthäus, B., Schumacher, J., Schall, P. *et al.* (2020). It is not
459 just a ‘trade-off’: indications for sink- and source-limitation to vegetative and regenerative
460 growth in an old-growth beech forest. *New Phytologist*, 226, 111–125.
- 461 Nussbaumer, A., Meusburger, K., Schmitt, M., Waldner, P., Gehrig, R., Haeni, M. *et al.* (2020).
462 Extreme summer heat and drought lead to early fruit abortion in european beech. *Scientific*
463 *Reports*, 10, 1–11.
- 464 Nussbaumer, A., Waldner, P., Apuhtin, V., Aytar, F., Benham, S., Bussotti, F. *et al.* (2018).
465 Impact of weather cues and resource dynamics on mast occurrence in the main forest tree
466 species in Europe. *Forest Ecology and Management*, 429, 336–350.
- 467 Owens, J.N. & Morris, S.J. (1998). Factors affecting seed and cone development in pacific silver
468 fir (*Abies amabilis*). *Canadian Journal of Forest Research*, 28, 1146–1163.
- 469 Pearse, I.S., Koenig, W.D. & Kelly, D. (2016). Mechanisms of mast seeding: resources, weather,
470 cues, and selection. *New Phytologist*, 212, 546–562.
- 471 Pearse, I.S., LaMontagne, J.M., Lordon, M., Hipp, A.L. & Koenig, W.D. (2020). Biogeography
472 and phylogeny of masting: do global patterns fit functional hypotheses? *New Phytologist*,
473 227, 1557–1567.
- 474 Pearse, I.S., Wion, A.P., Gonzalez, A.D. & Pesendorfer, M.B. (2021). Understanding mast
475 seeding for conservation and land management. *Philosophical Transactions of the Royal*
476 *Society B: Biological Sciences*, 376.
- 477 Pennekamp, F., Iles, A.C., Garland, J., Brennan, G., Brose, U., Gaedke, U. *et al.* (2019).
478 The intrinsic predictability of ecological time series and its potential to guide forecasting.
479 *Ecological Monographs*, 89, e01359.
- 480 Pesendorfer, M.B., Ascoli, D., Bogdziewicz, M., Hacket-Pain, A., Pearse, I.S. & Vacchiano,
481 G. (2021). The ecology and evolution of synchronized reproduction in long-lived plants.
482 *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376.
- 483 Pesendorfer, M.B., Bogdziewicz, M., Szymkowiak, J., Borowski, Z., Kantorowicz, W., Espelta,
484 J.M. *et al.* (2020). Investigating the relationship between climate, stand age, and temporal
485 trends in masting behavior of european forest trees. *Global Change Biology*, 26, 1654–1667.
- 486 Pesendorfer, M.B., Koenig, W.D., Pearse, I.S., Knops, J.M. & Funk, K.A. (2016). Individual
487 resource limitation combined with population-wide pollen availability drives masting in the
488 valley oak (*Quercus lobata*). *Journal of Ecology*, 104, 637–645.
- 489 Petchey, O.L., Pontarp, M., Massie, T.M., Kéfi, S., Ozgul, A., Weilenmann, M. *et al.* (2015).
490 The ecological forecast horizon, and examples of its uses and determinants. *Ecology Letters*,
491 18, 597–611.

- 492 Rammig, A., Fahse, L., Bebi, P. & Bugmann, H. (2007). Wind disturbance in mountain forests:
493 Simulating the impact of management strategies, seed supply, and ungulate browsing on forest
494 succession. *Forest Ecology and Management*, 242, 142–154.
- 495 Rogers, T.L., Johnson, B.J. & Munch, S.B. (2022). Chaos is not rare in natural ecosystems.
496 *Nature Ecology Evolution*, 6, 1005–1111.
- 497 Roncé, I.L., Dardevet, E., Venner, S., Schönbeck, L., Gessler, A., Chuine, I. *et al.* (2023).
498 Reproduction alternation in trees: testing the resource depletion hypothesis using experimental
499 fruit removal in quercus ilex. *Tree Physiology*.
- 500 Sala, A., Hopping, K., McIntire, E.J., Delzon, S. & Crone, E.E. (2012). Masting in whitebark
501 pine (*Pinus albicaulis*) depletes stored nutrients. *New Phytologist*, 196, 189–199.
- 502 Samarth, Kelly, D., Turnbull, M.H. & Jameson, P.E. (2020). Molecular control of masting: an
503 introduction to an epigenetic summer memory. *Annals of Botany*, 125, 851–858.
- 504 Samarth, Lee, R., Kelly, D., Turnbull, M.H., Macknight, R.C., Poole, A.M. *et al.* (2021).
505 Molecular control of the floral transition in the mast seeding plant *Celmisia lyallii* (Asteraceae).
506 *Molecular Ecology*, 30, 1846–1863.
- 507 Satake, A. & Iwasa, Y. (2000). Pollen coupling of forest trees: Forming synchronized and
508 periodic reproduction out of chaos. *Journal of Theoretical Biology*, 203, 63–84.
- 509 Satake, A. & Kelly, D. (2021). Studying the genetic basis of masting. *Philosophical Transactions
510 of the Royal Society B: Biological Sciences*, 376.
- 511 Schermer, E., Bel-Venner, M.C., Fouchet, D., Siberchicot, A., Boulanger, V., Caignard, T. *et al.*
512 (2019). Pollen limitation as a main driver of fruiting dynamics in oak populations. *Ecology
513 Letters*, 22, 98–107.
- 514 Schermer, E., Bel-Venner, M.C., Gaillard, J.M., Dray, S., Boulanger, V., Roncé, I.L. *et al.*
515 (2020). Flower phenology as a disruptor of the fruiting dynamics in temperate oak species.
516 *New Phytologist*, 225, 1181–1192.
- 517 Schmidt, K.A. & Ostfeld, R.S. (2008). Numerical and behavioral effects within a pulse-driven
518 system: Consequences for shared prey. *Ecology*, 89, 635–646.
- 519 Tattoni, C., Chianucci, F., Ciolli, M., Ferrara, C., Marchino, L., Zanni, M. *et al.* (2021). A
520 comparison of ground-based count methods for quantifying seed production in temperate
521 broadleaved tree species. *Annals of Forest Science*, 78, 11.
- 522 Touzot, L., Bel-Venner, M.C., Gamelon, M., Focardi, S., Boulanger, V., Débias, F. *et al.* (2018).
523 The ground plot counting method: A valid and reliable assessment tool for quantifying seed
524 production in temperate oak forests? *Forest Ecology and Management*, 430, 143–149.
- 525 Tseng, Y.T., Kawashima, S., Kobayashi, S., Takeuchi, S. & Nakamura, K. (2020). Forecasting
526 the seasonal pollen index by using a hidden markov model combining meteorological and
527 biological factors. *Science of The Total Environment*, 698, 134246.
- 528 Vacchiano, G., Ascoli, D., Berzaghi, F., Lucas-Borja, M.E., Caignard, T., Collalti, A. *et al.*
529 (2018). Reproducing reproduction: How to simulate mast seeding in forest models. *Ecological
530 Modelling*, 376, 40–53.

- 531 Vacchiano, G., Hacket-Pain, A., Turco, M., Motta, R., Maringer, J., Conedera, M. *et al.* (2017).
532 Spatial patterns and broad-scale weather cues of beech mast seeding in Europe. *New Phytol-*
533 *ogist*, 215, 595–608.
- 534 Wion, A.P., Pearse, I.S., Rodman, K.C., Veblen, T.T. & Redmond, M.D. (2023). Masting is
535 shaped by tree-level attributes and stand structure, more than climate, in a rocky mountain
536 conifer species. *Forest Ecology and Management*, 531, 120794.
- 537 Wion, A.P., Weisberg, P.J., Pearse, I.S. & Redmond, M.D. (2020). Aridity drives spatiotemporal
538 patterns of masting across the latitudinal range of a dryland conifer. *Ecography*, 43, 569–580.
- 539 Woolsey, J., Hanna, C., McEachern, K., Anderson, S. & Hartman, B.D. (2018). Regeneration
540 and expansion of *Quercus tomentella* (Island Oak) groves on Santa Rosa Island. *Western*
541 *North American Naturalist*, 78, 758–767.
- 542 Yates, L.A., Aandahl, Z., Richards, S.A. & Brook, B.W. (2023). Cross validation for model
543 selection: A review with examples from ecology. *Ecological Monographs*, 93, 1–24.
- 544 Zohner, C.M., Mo, L. & Renner, S.S. (2018). Global warming reduces leaf-out and flowering
545 synchrony among individuals. *eLife*, 7, e40214.

Acknowledgements

We thank Ian Pearse for his helpful comments on the earlier version of this manuscript, and Andreas Wion, Samuel Venner, and an anonymous reviewer for their constructive comments that helped us improve the text. VJ was supported by project No. 2021/43/P/NZ8/01209 co-funded by the Polish National Science Centre and the EU H2020 research and innovation programme under the MSCA GA No. 945339. For the purpose of Open Access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission. MB was supported by the European Union (ERC, ForestFuture, 101039066). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. MBP and IO were supported by the Waldfonds project FORSEE, funded by the Austrian Federal Ministry of Agriculture, Regions and Tourism. The development of MASTREE+ was supported by the UK Natural Environment Research Council grant no. NE/S007857. AHP received funding from Defra through a UKRI Future of UK Treescapes programme Fellowship. We acknowledge the organizers and contributors of the British Ecological Society-Royal Meteorological Society symposium "Climate Science for Ecological Forecasting", London, May 2022.

Author contributions

All authors designed and developed the study. Analysis was performed by VJ. MB and VJ led the writing of the manuscript with contributions from AHP, IO and MBP.

Competing interests

None declared.

Data availability statement

Data used are available at : <https://doi.org/10.17605/OSF.IO/RKN8A>.

574 **New Phytologist Supporting Information**

575 **Article title:** Forecasting seed production in perennial plants: identifying challenges and chart-
576 ing a path forward

577 **Authors:** Valentin Journé, Andrew Hacket-Pain, Iris Oberklammer, Mario B. Pesendorfer,
578 Michał Bogdziewicz

579 **Article acceptance date:** 12-04-2023

580

581 **Methods S1**

582 **Seed production data**

583 We used data from MASTREE+ database (Hacket-Pain *et al.*, 2022) and extracted seed produc-
584 tion data for European beech *Fagus sylvatica*. We only used records measured on a continuous
585 scale that covered more than 8 years of observations after 1980, and excluded records of annual
586 flowering, pollen production, and tree-ring-based mast year reconstructions (Fig. A1a) ($n = 94$
587 MASTREE+ sites; 1,929 site-years). We scaled seed production for each site to between 0 and
588 1 by dividing all values by the highest recorded value in the time series (Fig. A1b). This was
589 done to accommodate among-site variation in methods of seed production monitoring among
590 various datasets included in MASTREE+.

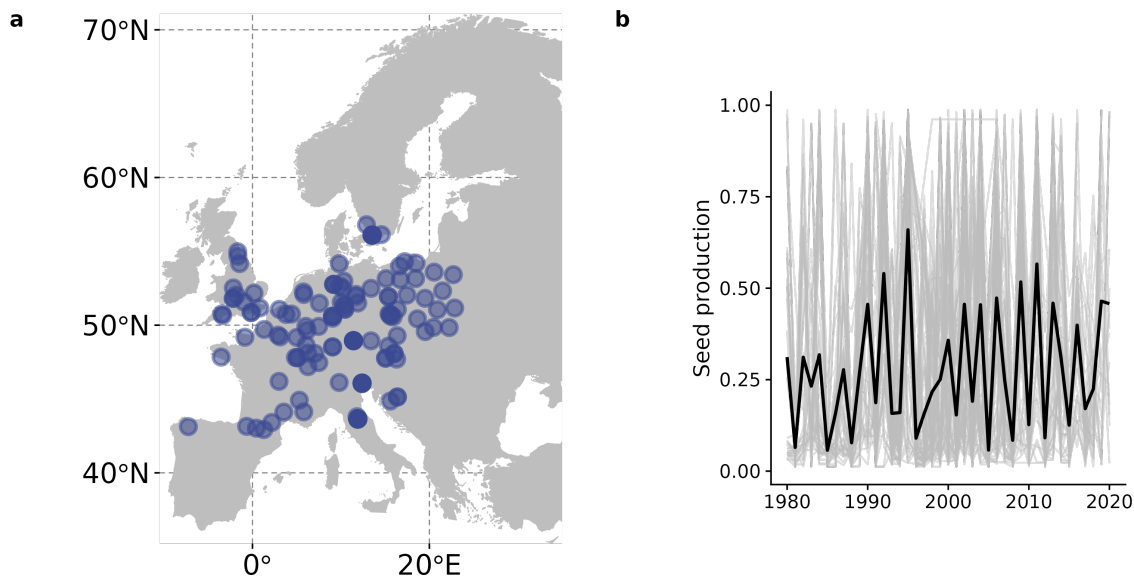


Figure A1: a) Map of sites used in the analysis ($n=94$). b) Seed production patterns for each site are shown in grey with an average value from all sites in black.

591 **Models : ΔT , *foreMast*, and sequential model**

592 We used three different models to predict masting in beech: ΔT (Kelly *et al.*, 2013), *foreMast*
593 (Chiavetta & Marzini, 2021), and the sequential model. The ΔT model is based on the difference
594 in temperature between two past consecutive summers ($T1$ and $T2$) (Kelly *et al.*, 2013). The

595 *foreMast* model is based on an R package that predicts the probability of seed production in a
 596 given year (Chiavetta & Marzini, 2021). This model includes both previous ($T1$ and $T2$) summer
 597 temperature and summer precipitation but also an auto-correlation term that helps to avoid two
 598 consecutive high-seed production years. The sequential model includes several weather variables
 599 known to affect seed production in beech, from pollination to fruit maturation, and the previous
 600 year's seed production: past summer temperature ($T1$ and $T2$), spring temperature of the year of
 601 seed fall ($T0$), late frost and summer moisture deficit of the year of seed fall; all defined below.

602 The *foreMast* algorithm provides masting forecasts for any given set of coordinates, while
 603 the coefficients for the ΔT and sequential models were estimated by fitting separate models to
 604 each time series (i.e. each unique site). We have fitted site-specific models, instead of for
 605 example a hierarchical model, as we were interested in site-specific coefficients for predictions.
 606 We used beta regression to fit the ΔT and sequential model (Grün *et al.*, 2012; Brooks *et al.*,
 607 2017). As the beta distribution allows only values above zero and below one, and re-scaled our
 608 data for each site using the equation (Smithson & Verkuilen, 2006; Grün *et al.*, 2012):

$$Y_s = \frac{(y_s * (n_s - 1) + 0.5)}{n_s} \quad (1)$$

609 the y_s is the observation value of seed production for the site s (previously scaled to $0 \leq y_s$
 610 ≤ 1) and n_s is the number of observation per site. This method involves the addition of minute
 611 values to observations that possess a value of 0, while simultaneously subtracting negligible
 612 values from observations that possess a value of 1.

613 For ΔT and the sequential model, we extracted daily minimal, maximal, and average tem-
 614 perature, and precipitation used for the moisture deficit, for each site using EOS-obs (v22.0e,
 615 Cornes *et al.* 2018, $0.1^\circ \times 0.1^\circ$). Summer temperature was defined as the average of June-July-
 616 August max temperature. Spring temperature was April-May mean average temperature. We
 617 defined late spring frost based on cumulative growing degree days calculated for each year and
 618 MASTREE+ site as the sum of daily mean air temperature above 5°C from the 1st of January to
 619 the date of the last frost day ($< -2^\circ\text{C}$) (Vitasse *et al.*, 2018). A high value of cumulative degree
 620 days indicates a high risk of frost injury (Vitasse *et al.*, 2018). The summer moisture deficit
 621 was defined as $P - PET$, with P for precipitation and PET for evapotranspiration, summed from
 622 June 1st to August 31st. Evapotranspiration (PET) was calculated with the Thornthwaite equa-
 623 tion (Thornthwaite, 1948), by using the R package SPEI (Beguería & Vicente-Serrano, 2017).
 624 Distribution and correlation between climate variables are reported in Figure S8. Covariates
 625 included in the sequential model were standardized and centered before being included in the
 626 model.

627 The *foreMast* package includes a function to extract climate from each location starting
 628 from 1981 based on ERA5-Land (Chiavetta & Marzini, 2021). Both monthly-year average
 629 temperature and precipitation, coming from EOS-obs and *foreMast* climate extraction present
 630 robust correlation (Figure S9).

631 To evaluate the effects of focal predictors on the sequential model predictive power, we
 632 fitted several sub-models. This included two approaches: in one, we excluded one variable at
 633 a time (e.g. spring temperature in year $T0$), refitted the model, and evaluated its performance.
 634 In the second, we added variables in their phenological order of occurrence, starting with a
 635 model with only summer $T2$ (i.e. only one predictor), next added model summer $T1$ (two
 636 predictors), previous year seed production (tree predictors), late spring frost (four predictors),
 637 spring temperature (five predictors), and lastly summer moisture deficit (six predictors).

638 We assessed the ability of models to re-create past seed production by using two main
 639 metrics: the *R-squared* (R^2) and the *Root Mean Squared Error* ($RMSE$). The objective is to

640 compare predictions of yearly seed production to the observed yearly seed production. The R^2
641 is the square of the correlation, which gives the proportion of the variation in the dependent
642 variable that is predictable from the independent variables. An R^2 of 1 means that predictions
643 perfectly match the observations. The $RMSE$ provides the standard deviation of the prediction;
644 smaller values indicate better model performance. Models were fitted on the site level, site-
645 specific parameter coefficients were used to back-cast seed production, and R^2 and $RMSE$ were
646 computed for each site.

647 We also conducted cross-validation on the sequential model by using a cross-validation
648 procedure (Arlot & Celisse, 2010). We randomly sampled 70% of the data in each time series
649 to train the model and validate it based on the remaining 30% (by computing R^2 and $RMSE$ for
650 each site). We repeated this random sampling 10 times for each site and then computed the R^2
651 for each site.

652 References

- 653 Arlot, S. & Celisse, A. (2010). A survey of cross-validation procedures for model selection.
654 *Statistics Surveys*, 4, 40–79.
- 655 Beguería, S. & Vicente-Serrano, S.M. (2017). *SPEI: Calculation of the Standardised*
656 *Precipitation-Evapotranspiration Index*. R package version 1.7.
- 657 Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A.
658 *et al.* (2017). glmmTMB balances speed and flexibility among packages for zero-inflated
659 generalized linear mixed modeling. *The R Journal*, 9, 378–400.
- 660 Chiavetta, U. & Marzini, S. (2021). foreMast: an R package for predicting beech (*Fagus sylvatica*
661 L.) masting events in European countries. *Annals of Forest Science*, 78, 1–10.
- 662 Cornes, R.C., van der Schrier, G., van den Besselaar, E.J. & Jones, P.D. (2018). An ensem-
663 ble version of the E-OBS temperature and precipitation data sets. *Journal of Geophysical*
664 *Research: Atmospheres*, 123, 9391–9409.
- 665 Grün, B., Kosmidis, I. & Zeileis, A. (2012). Extended beta regression in R: Shaken, stirred,
666 mixed, and partitioned. *Journal of Statistical Software*, 48, 1–25.
- 667 Hacket-Pain, A., Foest, J.J., Pearse, I.S., LaMontagne, J.M., Koenig, W.D., Vacchiano, G. *et al.*
668 (2022). MASTREE+: Time-series of plant reproductive effort from six continents. *Global*
669 *Change Biology*, 28, 3066–3082.
- 670 Kelly, D., Geldenhuis, A., James, A., Penelope Holland, E., Plank, M.J., Brockie, R.E. *et al.*
671 (2013). Of mast and mean: Differential-temperature cue makes mast seeding insensitive to
672 climate change. *Ecology Letters*, 16, 90–98.
- 673 Smithson, M. & Verkuilen, J. (2006). A better lemon squeezer? Maximum-likelihood regression
674 with beta-distributed dependent variables. *Psychological Methods*, 11, 54–71.
- 675 Thornthwaite, C. (1948). An approach toward a rational classification of climate. *Geographical*
676 *Review*, 38, 55–94.
- 677 Vitasse, Y., Schneider, L., Rixen, C., Christen, D. & Rebetez, M. (2018). Increase in the risk of
678 exposure of forest and fruit trees to spring frosts at higher elevations in Switzerland over the
679 last four decades. *Agricultural and Forest Meteorology*, 248, 60–69.

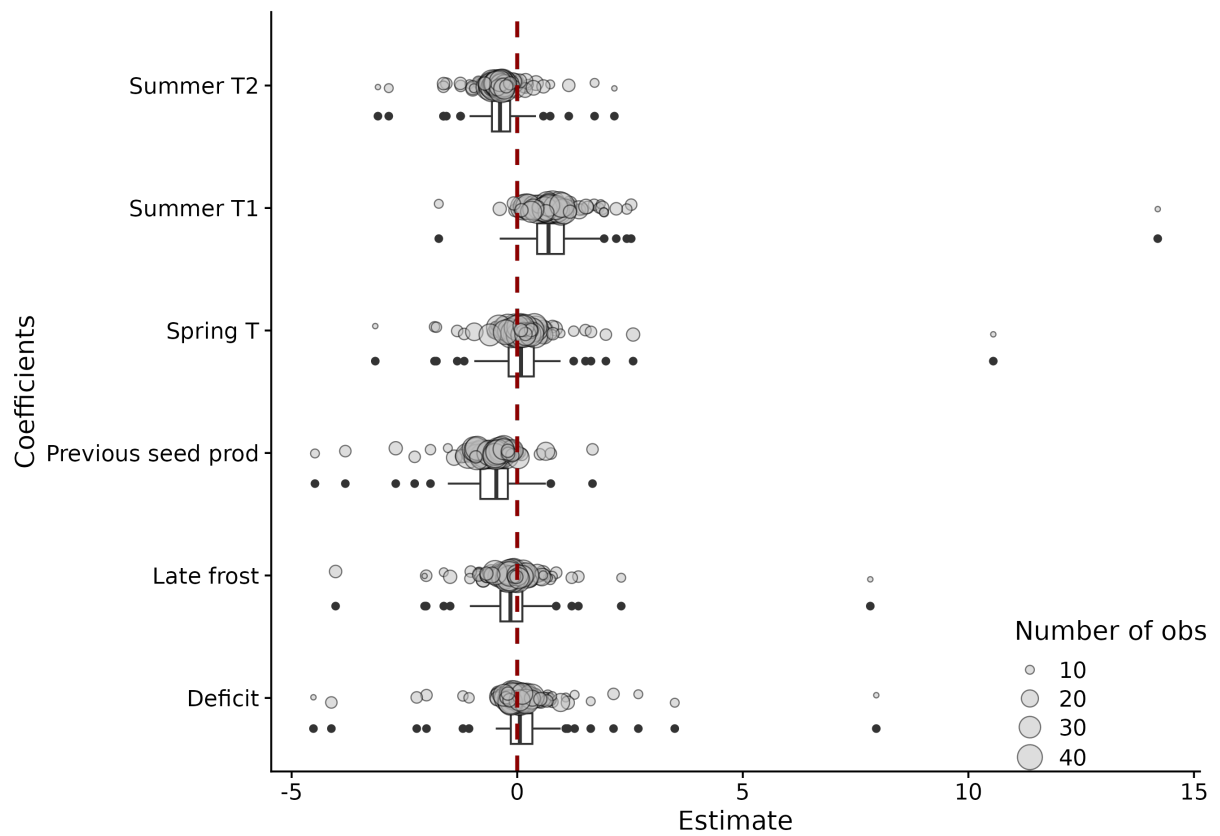


Figure S1: Standardized coefficients from the sequential model coming from MASTREE+ sites. Coefficient values have been extracted for each covariate and from each site where the beta regression models have been run. Grey dots represent the coefficient values of the site. The dot size above the boxplots is scaled to the number of observations that have been used to run beta regression models at the focal site.

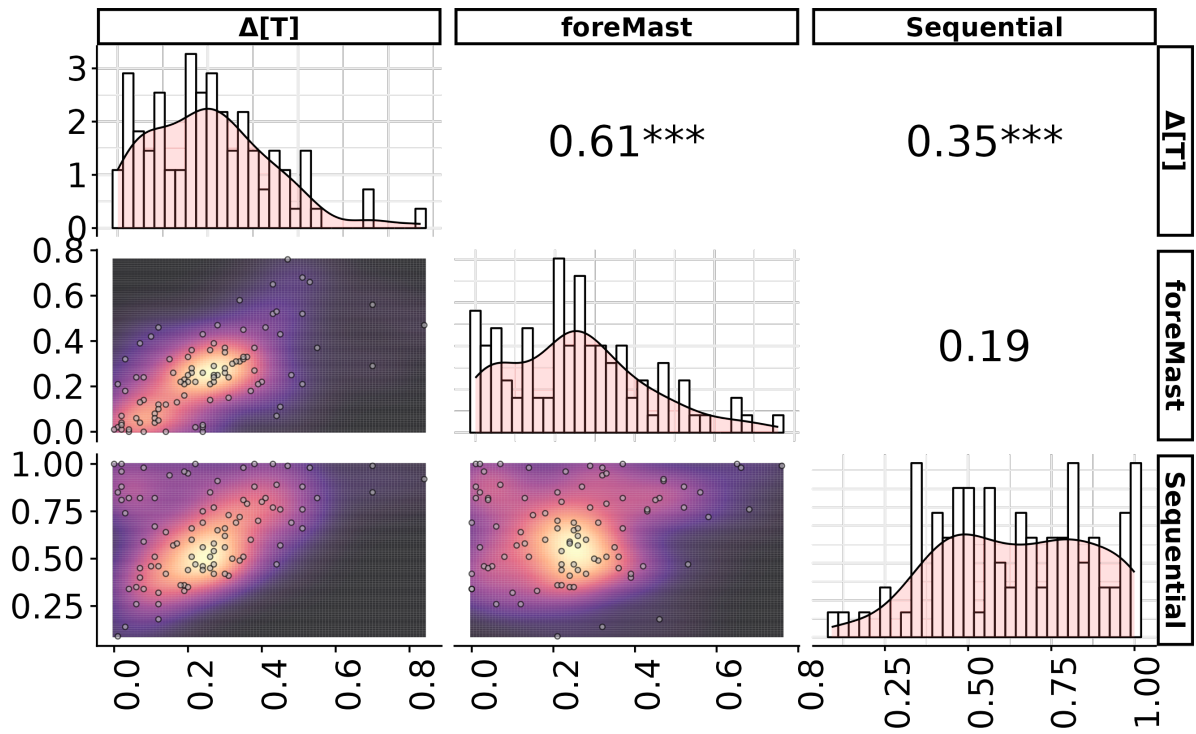


Figure S2: R^2 comparisons among the three models: scatter plots of R^2 , R^2 distributions, and their correlation. For correlations, we used Pearson correlation with *** for P -values < 0.001 , ** for 0.001, * for 0.01. The range of x-axis (R^2) differs between the three models. The lower panel presents kernel density estimation of R^2 , from dark to yellow for higher density, and each dots represent R^2 of each sites.

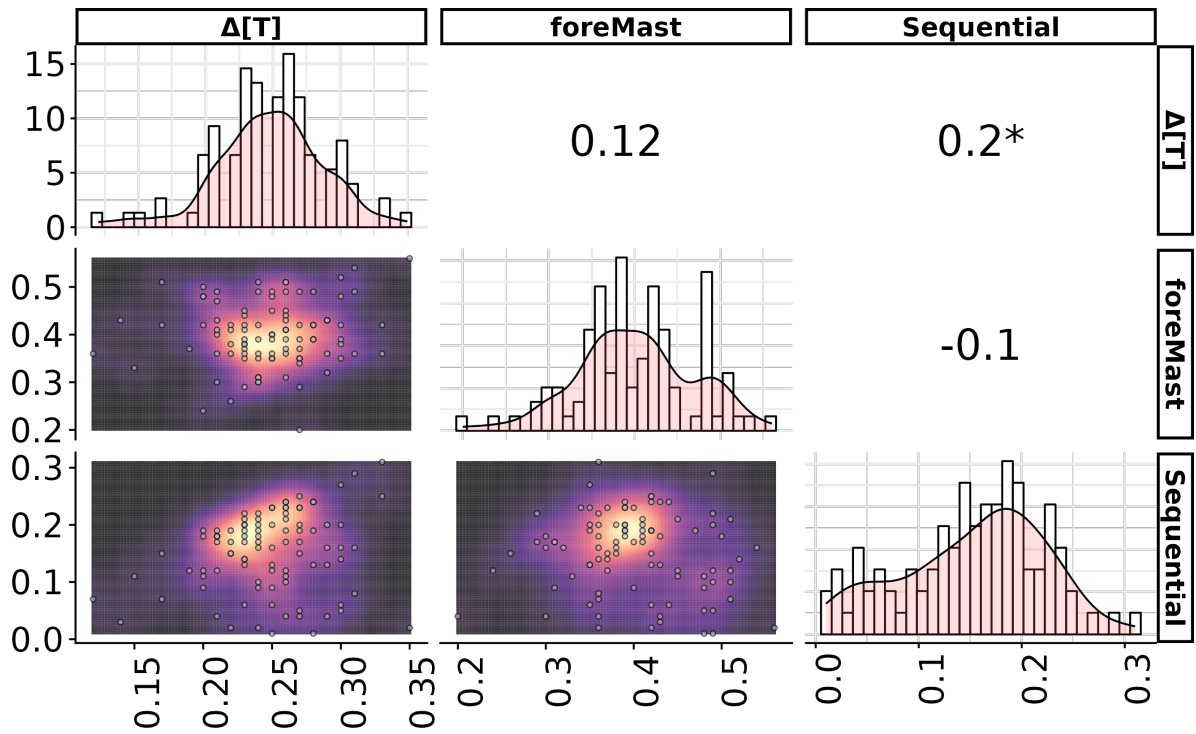


Figure S3: *RMSE* comparisons among the three models: scatter plots, *RMSE* distributions, and their correlation. For correlations, we used Pearson correlation with *** for P -values < 0.001 , ** for 0.001, * for 0.01. The range of x-axis metric (*RMSE*) differs between the three models. The lower panel presents kernel density estimation of *RMSE*, from dark to yellow for higher density, and where each dots represent *RMSE* of each site.

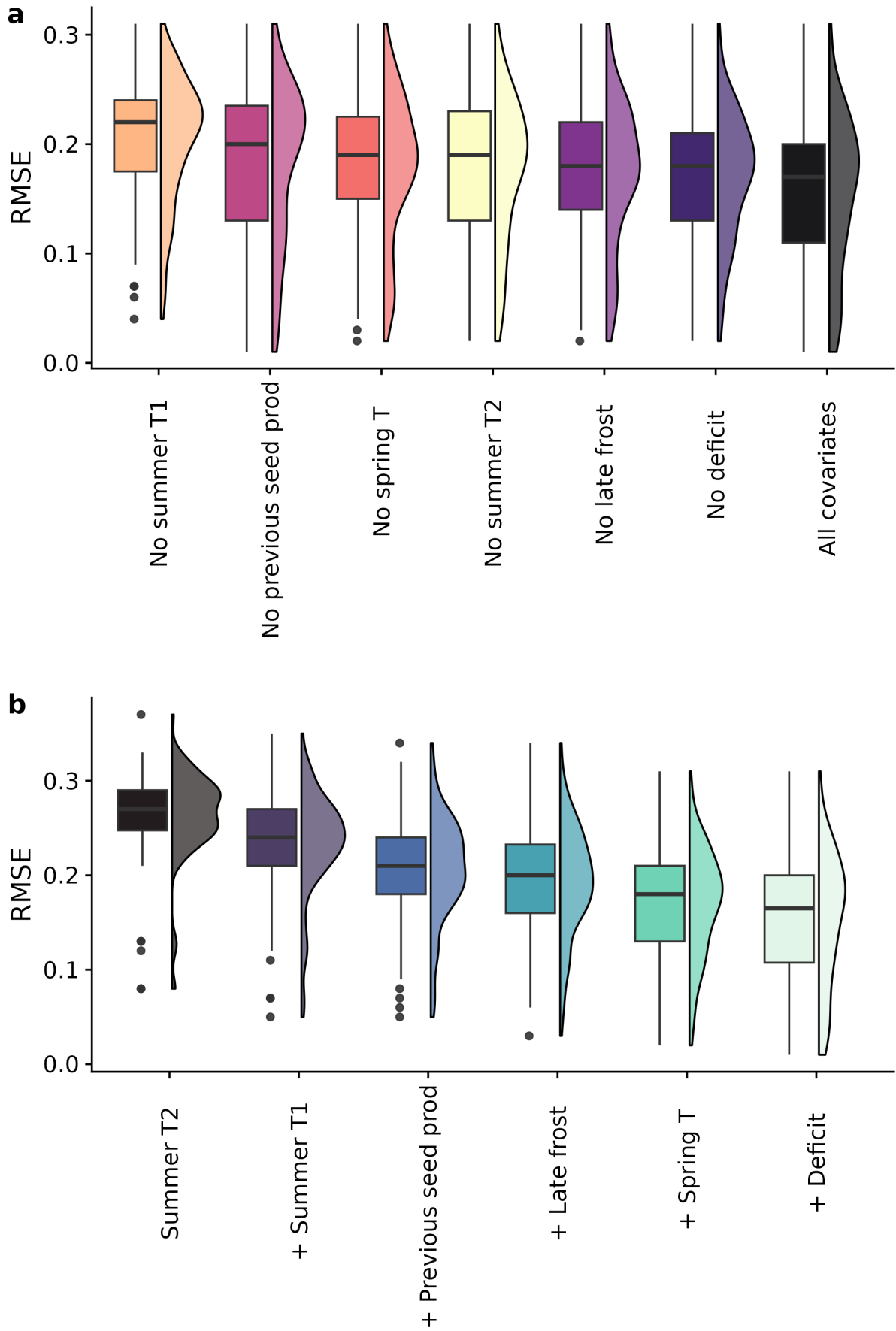


Figure S4: Effects of focal predictors on the sequential model performance evaluated with *RMSE*. a) Each boxplot shows sequential model *RMSE* with one predictor removed, b) predictors added in phenological order.

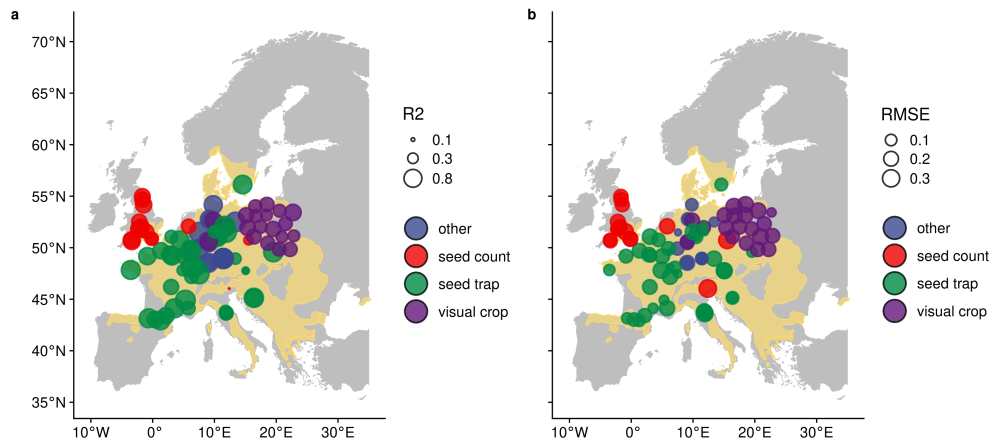


Figure S5: Goodness-of-fit of the sequential model presented with a) R^2 and b) $RMSE$. The yellow background represents the species distribution of *Fagus sylvatica* extracted from EUFORGEN (<https://www.euforgen.org/species/>). The dots represent each site, color-coded by the collection method. The size of the dots varies according to goodness-of-fit. The "other" category follows MASTREE+ and gathers collection methods that did not fit any other categories presented.

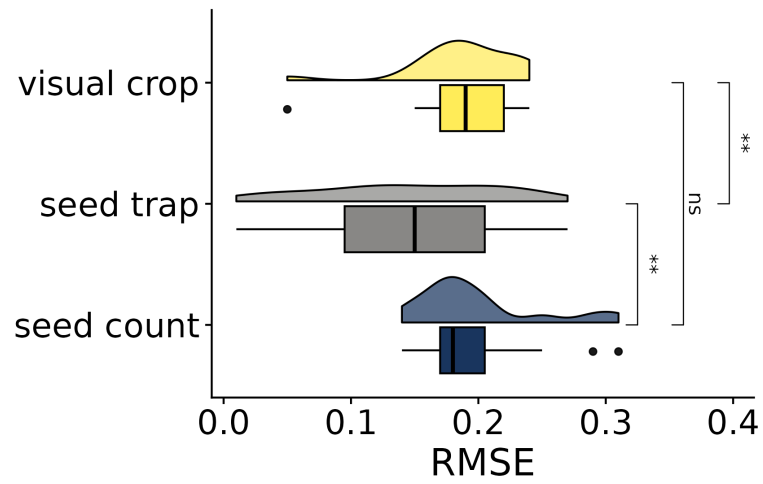


Figure S6: Effects of seed production estimation method on the sequential model *RMSE*. The methods are seed count (ground counting within a given time frame), monitoring with seed traps, and visual crop assessment. The asterisks represent significant differences between each group tested with a t-test (**** for $P < 0.0001$, *** for $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, and n.s. for $P > 0.05$)

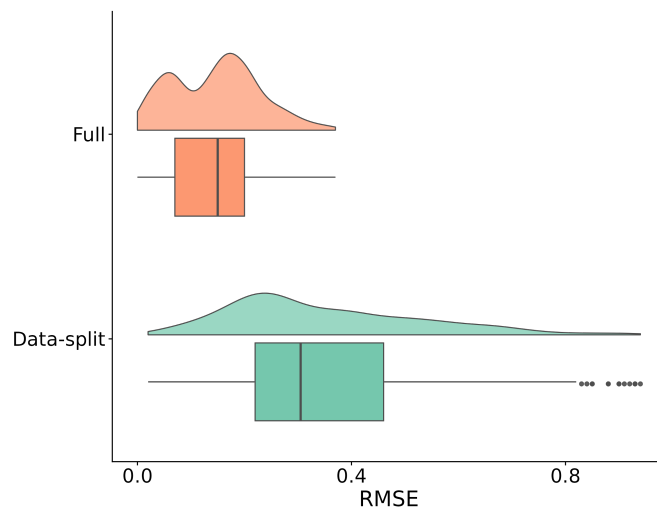


Figure S7: Boxplot of the $RMSE$ from the model trained on 70% of the dataset or trained on the full dataset.

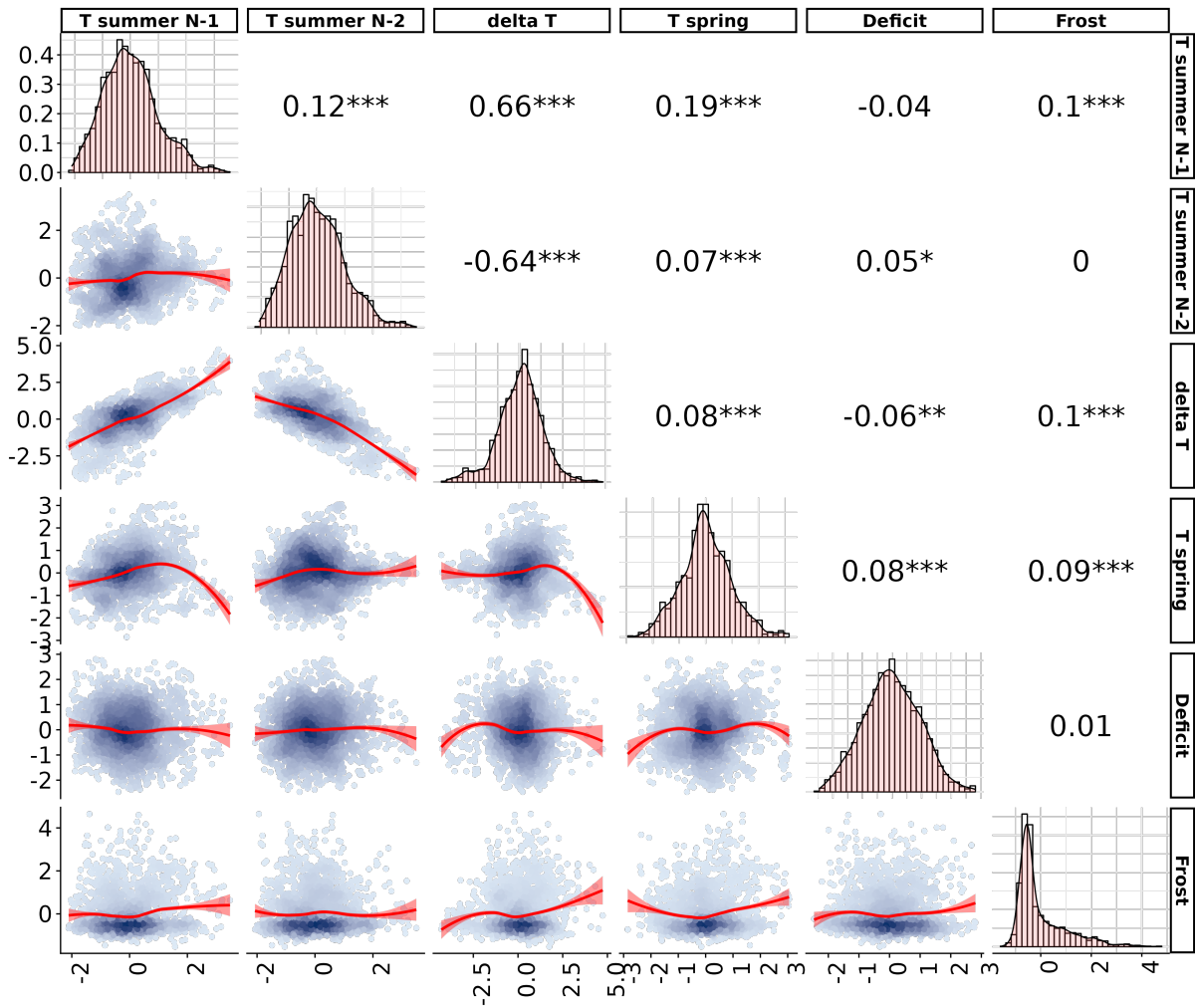


Figure S8: Bivariate relationships and distribution of climatic variables. The upper panels show correlations, and the lower panels bivariate relationships using the loess function. Dots represent each site-yearly observations. For correlation between scaled climate variables, we used Pearson correlation with *** for P -values <0.001 , ** for <0.001 , * for <0.01

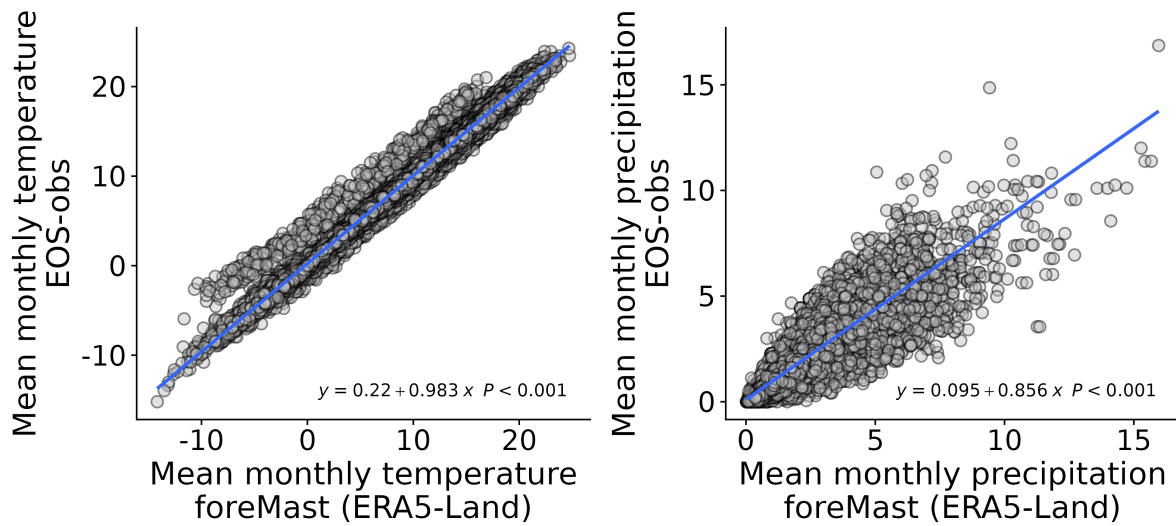


Figure S9: Correlation between climate obtained from EOS-obs and *foreMast* package (based on ERA5-Land). a) Correlation between average monthly temperature and b) Correlation between average monthly precipitation. Each dot represents a value for a month-year from our sites. Linear model summary is reported with slope, intercept and *P*-values at the bottom right.