Comparing statistical methods for detecting climatic drivers of mast seeding

—

Valentin Journé*[1,2], Emily G. Simmonds[3], Maciej K. Barczyk[2], Michał Bogdziewicz[2]

[1]Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan
[2]Forest Biology Center, Faculty of Biology, Institute of Environmental Biology, Adam Mickiewicz University, Poznan, Poland.
[3]Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, UK.
*corresponding author: journe.valentin@gmail.com

Abstract

Understanding the drivers of mast seeding is critical for predicting reproductive dynamics in perennial plants. Here, we evaluate the performance of four statistical methods for identifying weather-associated drivers of annual seed production, i.e, weather cues: climate sensitivity profile, P-spline regression, sliding window analysis, and peak signal detection. Using long-term seed production data from 50 European beech (*Fagus sylvatica*) populations and temperature records, we assessed each method's ability to detect a benchmark window around the summer solstice. All methods successfully identified biologically meaningful windows, but their performance varied with data quality, signal strength, and sample size. Sliding window and climate sensitivity profile methods showed the best balance of accuracy and robustness, while peak signal detection had lower consistency. Cue identification was more reliable with at least 20 years of data, and predictive accuracy was highest when models were based on seed trap data. A simulation study showed method-specific sensitivity to signal strength, with the sliding window performing best. Our findings provide a means to improve masting forecasts through a practical guide for selecting appropriate cue identification methods under varying data constraints.

keywords: phenology | seed production | weather | climate change

Data and code availability statement

Data are archived on the OSF Repository at the following link: https://osf.io/u23vy/. The case study code is accessible on Github https://github.com/ValentinJourne/weatheRcues/tree/main/Application_MASTREE. A R vignette tutorial is available at https://valentinjourne.github.io/weatheRcues/articles/weatheRcues.html.

Author contributions

VJ, EGS and MBo conceived and designed the study. VJ led the analysis and wrote the code, with additions from EGS and MBa. MBo and VJ led the writing of the manuscript with contributions from all authors. All authors revised the paper.

correspondence:

Journé Valentin*: journe.valentin@gmail.com

*Department of Biology, Faculty of Science, Kyushu University, Fukuoka, Japan.

Introduction

Mast seeding, or masting, is synchronous and highly variable reproduction among years by a population of perennial plants (Kelly, 1994; Pearse *et al.*, 2016). Masting increases pollination efficiency and reduces seed predation, enhancing reproductive success (Kelly *et al.*, 2001; Zwolak *et al.*, 2022; Bogdziewicz *et al.*, 2024a). Moreover, interannual variation in seed production generates resource pulses that shape ecosystem functioning by influencing seed consumer populations and, in turn, their predators, parasites, and scavengers (Clark *et al.*, 2019; Maag *et al.*, 2024; Widick *et al.*, 2025). Masting also influences tree growth, defense investment, nutrient cycling, and the abundance of mycorrhizal fungi (Hacket-Pain *et al.*, 2018; Redmond *et al.*, 2019; Michaud *et al.*, 2024). Thus, ecosystem management and conservation of plants and animals require a comprehensive understanding of the drivers of masting (Pearse *et al.*, 2021). Among these drivers, weather cues, i.e., weather variation that regulates processes such as flower initiation, pollination success, and fruit maturation, play a major role (Kelly *et al.*, 2013; Bogdziewicz *et al.*, 2025). Given the large variations in weather cues among species and populations (Bogdziewicz *et al.*, 2019; Koenig *et al.*, 2020; Fleurot *et al.*, 2023), and historic data scarcity (Koenig, 2021), the relationships between seeding and weather are largely unidentified or rely on, often somewhat arbitrary, a priori selection (Crone & Rapp, 2014). Yet, identification of weather cues is a key step in analyses aimed at masting forecasting important in conservation in management (Wion *et al.*, 2025; Oberklamme *et al.*, 2025), advancing understanding of masting biology (Journé *et al.*, 2024; Szymkowiak *et al.*, 2024), and improving predictions of the effects of climate change on masting dynamics (LaMontagne *et al.*, 2021; Bogdziewicz *et al.*, 2024b). Here, we evaluate four statistical methods designed to identify periods of strongest correlation between weather variation and biological responses to detect weather cues of masting.

Masting plants have evolved hypersensitivity to weather variation, a trait that amplifies interannual variation in seed production relative to the variation in weather fluctuations (Janzen, 1971; Kelly *et al.*, 2013). This hypersensitivity enables plant populations to synchronize reproductive efforts by collectively delaying reproduction under unfavorable conditions and initiating mass seeding events when conditions improve (Abe *et al.*, 2016; Schermer *et al.*, 2020; Ascoli *et al.*, 2020). Weather variation influences seed production across multiple stages of the fruit maturation cycle, including flower initiation, pollination, and fruit maturation (Pearse *et al.*, 2016; Bogdziewicz *et al.*, 2025). Specific weather cues differ among species and

populations. For example, increased flower bud initiation may follow hot summers, whereas reproduction may largely fail when spring weather hinders effective pollen transfer (Koenig *et al.*, 2015; Fleurot *et al.*, 2024; Journé *et al.*, 2024). Some species exhibit conserved cues and mechanisms; European beech (*Fagus sylvatica*), for example, relies on summer temperatures across its range (Journé *et al.*, 2024). In contrast, sessile oaks (*Quercus petraea*) display spatial variation in masting regulation, with spring temperatures governing pollination success and fruit set in semi-continental climates and flower number being a primary determinant in oceanic climates (Fleurot *et al.*, 2023).

Finding the best weather correlate for seed production for a particular population is important. In global change biology, shifts in weather cue frequency associated with climate change can alter masting patterns (Shibata *et al.*, 2020; Foest *et al.*, 2024), with effects varying according to whether reproduction associates with warm or cold temperatures (Bogdziewicz *et al.*, 2024b). When reproduction relies on warm-associated cues, warming increases cue frequency, leading to more frequent but smaller flower crops and reduced synchrony among individuals (Bogdziewicz *et al.*, 2021; Foest *et al.*, 2024). Such changes diminish the benefits of masting and can result in dramatic reductions in viable seed production (Bogdziewicz *et al.*, 2020b, 2023b). What is more, the increased frequency of reproduction leads to growth reduction (Hacket-Pain *et al.*, 2025). In contrast, in plants that rely on cues associated with low temperatures, warming leads to fewer reproductive attempts, resulting in prolonged periods of seed failure (Numata *et al.*, 2022; Yukich-Clendon *et al.*, 2023). Thus, determining the timing of cues and establishing the direction of the relationship between weather cues and reproduction are essential steps in predicting the impact of climate change on masting patterns and the reproductive capacity of plants (Bogdziewicz *et al.*, 2024b). Moreover, forecasts of seed production depend on a solid understanding of the links between weather and reproductive output (Journé *et al.*, 2023; Wion *et al.*, 2025; Oberklamme *et al.*, 2025). Thus, accurate identification of seed production-weather correlates reinforces effective management and conservation in ecosystems that include masting species (Pearse *et al.*, 2021). Finally, detecting correlation signals guides experimental manipulations by informing both the timing of interventions and the choice of environmental variables, such as temperature or precipitation (Pérez-Ramos *et al.*, 2010; Samarth *et al.*, 2021). Reliable methods of weather cue identification from increasingly available long-term data (Hacket-Pain *et al.*, 2022; Foest *et al.*, 2024) are now needed.

In this study, we compared four approaches to investigating relationships between seed

production and weather variation: climate-sensitive profiling, P-spline regression, sliding window analysis, and peak signal identification (Table 1) (Roberts, 2008; Simmonds *et al.*, 2019; Bailey & van de Pol, 2016; Lee *et al.*, 2024). Due to the limitations of observational studies and the logistical challenges of experimentally manipulating environmental signals in trees (Bogdziewicz *et al.*, 2020a), the true causal relationships between weather variation and seed production remain elusive (Pearse *et al.*, 2014; Pesendorfer *et al.*, 2021). The lack of an unequivocal reference for these relationships poses a significant challenge for validating statistical methods. To address this limitation, we used the well-documented relationship in European beech (*Fagus sylvatica*) as a benchmark. European beech exhibits robust negative correlations between seed production and June–July temperatures two years prior to seed fall, and positive correlations one year prior (Piovesan & Adams, 2011; Vacchiano *et al.*, 2017; Nussbaumer *et al.*, 2018; Bogdziewicz *et al.*, 2023a). The timing of these effects is consistent across the species' range, as the window is anchored to the summer solstice (Journé *et al.*, 2024). We assessed how each method detects these known patterns under varying conditions of data quality and sample size, and conducted a simulated case study in which the strength of the weather–seed production relationship was altered. By doing so, we test whether the focal methods can detect the benchmark cue without any prior assumptions, and we hope that this analysis can guide future applications of weather cue identification. We foresee that this aspect will become increasingly important for advancing understanding of masting dynamics.

Table 1: Summary of weather cue identification methods used in this study to determine the timing of weather cues. That is, to identify the specific time window when plants are most sensitive to variation in a given weather variable. The paper is supplemented with an R compendium that can facilitate the implementation of the focal methods.

Method	Summary	Reference
Climate sensitivity profile	This method includes running a linear regression between annual seed production and weather variables for each day, extracting the slope and R^2 values. To smooth these relationships over time, two generalized additive models (GAMs) are fitted using β_{days} (slope) and R^2_{days}, as responses. The weather cue window is the consecutive days during which the slope and R^2 values exceed the lower and upper quantiles (2.5th and 97.5th percentiles obtained from all days).	Thackeray *et al.* (2016); Simmonds *et al.* (2019)
P-spline regression	Similar to the climate sensitivity profile approach, but it differs in that partial coefficients are smoothed by applying a penalty to differences between consecutive days. Can handle multiple explanatory variables (e.g., all individual daily measurements) in a single analysis by using a data reduction step to address high dimensionality. This is done by constructing a B-spline function, consisting of piecewise polynomial curves connected at predefined knots. By combining B-splines with a difference penalty, P-splines (penalized B-splines) are created, preventing overfitting by penalizing excessive variation between the B-splines.	Roberts (2008); Roberts *et al.* (2015)
Peak signal detection	In the peak signal detection approach, the weather time series is systematically shifted by one day relative to annual seed production, and at each lag, the regression between the two is calculated. The lag that produces the highest correlation is interpreted as the time at which the weather variable most strongly relates to seed production. This method provides a direct measure of the optimal lag, although it relies on identifying a single peak in the correlation function. The signal strength for each day is determined by multiplying the coefficient of determination (R^2) by the slope. A peak signal detection is then used to identify the specific days with the strongest influence on seed production.	Brakel (2014); Lee *et al.* (2024)
Sliding window analysis	This approach tests a range of candidate time windows over which environmental data (e.g., temperature) are aggregated. For each window, defined by its start and end days, a summary statistic (e.g., mean) is calculated. A regression model is then fitted to assess how this aggregated weather variable explains variation in annual seed production. Models are compared using AIC, and the window with the best performance (lowest AIC) is identified as the optimal period of environmental influence.	van de Pol *et al.* (2016); Bailey & van de Pol (2016)

Methods

Seed production data

Our analysis is based on MASTREE+, a database of annual records of population-level reproductive effort in perennial plants from all vegetated continents (Hacket-Pain *et al.*, 2022; Foest *et al.*, 2024). We extracted data on European beech (*Fagus sylvatica*) and restricted the analysis to time series that included at least 20 years of records, observed after 1952 and before 2021, the latter done to match the weather dataset. We excluded flower and pollen counts, and ordinal records of seed production. In total, we used 50 time series, with a median length of 43 years (max = 67 years, Figure 1). We log-transformed annual seed production for each population to normalize data and ensure compatibility with the different cue-identification methods, all of which assume a Gaussian distribution.

In the MASTREE+ dataset, annual seed production is estimated at the population level using various methods, including seed counts – ground counting within a certain time frame (Foest *et al.*, 2025a), seed traps (Bajocco *et al.*, 2021), and visual crop assessment. Visual assessments are often used over large areas by foresters, for example, by the Polish State Forests (Pesendorfer *et al.*, 2020). These methods may differ in how well they capture among-year variation in seed production (Foest *et al.*, 2025a). Thus, we have assessed whether the performance of focal methods of weather cue window detection varies across methods of seed monitoring. Out of the 50 populations used in our analysis, 14 have annual records of seed production based on seed count, 17 used seed traps, and 19 used visual crop assessment.

Climate data

We extracted daily average temperature data for each site from the corresponding 0.1° grid cell of the E-OBS dataset (Cornes *et al.*, 2018) (version 28.0). The temperature was available from 1950 to 2023. We standardized the average temperature for each time series to ensure comparability and facilitate a simulation case study.
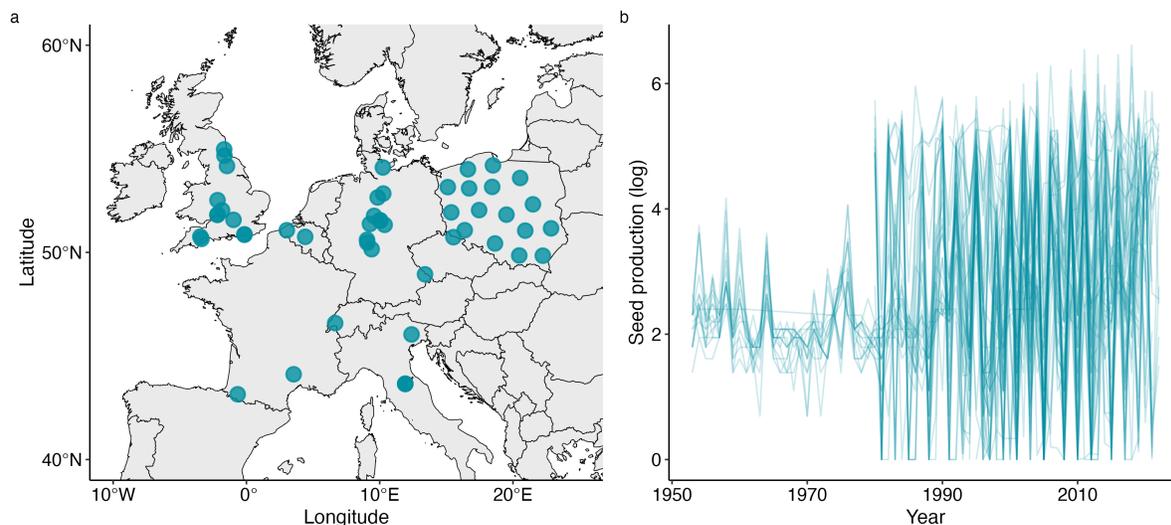
Figure 1: a) Map of the 50 European beech populations included in the study (minimum time series length: 20 years, median: 43 years). Each dot represents a single population. b) Temporal dynamics of seed production. Each line shows one population, with seed production values log-transformed (+1).

Description of the weather cue identification methods

For each of the four methods, we established the starting reference point as November 1st of the year associated with seed fall. We defined a time range extending from this reference date back to 600 days prior. This duration was chosen to include potential influences from summer temperature cues in both the first and second year preceding seed fall (Vacchiano *et al.*, 2017; Journé *et al.*, 2024).

Sliding window

The absolute positioning of the window opening and closing is defined by setting an origin point, from which the window moves backward in time (here, reference day, 1st November). In this approach, the window extends up to 600 days into the past, with a step length of one day. Additionally, the window length varies, ranging from a single day to a maximum of 600 days. When testing different windows, an aggregation method must be specified, for

instance, using the mean, maximum, or minimum of all daily weather values within a window. We used the climwin R package (version 1.2.3) (Bailey & van de Pol, 2016) to run the sliding window, and chose the mean temperature aggregation within the window. Climwin reports the best model based on the Akaike Information Criterion corrected for small sample sizes (AICc) (Bailey & van de Pol, 2016).

Peak signal detection

We regressed daily mean temperature against annual seed production for each day of the year, starting from the reference date until 600 days before, using a linear model. We extracted the slope and the model coefficient of determination (R^2) from each regression. Model strength was then determined by calculating the product of the slope and coefficient of determination ($\beta * R^2$), which measures the model's explanatory power. To detect peaks and valleys of model strength within the time series, we use a robust peak detection algorithm based on a z-score thresholding approach. The algorithm uses a rolling window, defined by a lag parameter, to calculate both the mean and standard deviation of model explanatory power. At each step, the algorithm flags a "signal" if the model's explanatory power for a day deviates from the local moving mean by more than a predefined number of standard deviations (the threshold) (Brakel, 2014). In our case, we used a lag of 100 days and a threshold of 3 standard deviations. The algorithm includes an influence parameter (here set at 0), which controls how much these identified signals affect future calculations of the moving mean and standard deviation, in order to prevent future bias in signal detection. Since this method can identify multiple potential windows due to multiple peaks per time series, we retained only one window for each time series based on the highest R^2.

Climate sensitivity profile

In this approach, the daily mean temperature is regressed against annual seed production for each day of the year. In our case, we started with the mean temperature on the reference day and regressed it against the seed production of the focal year using a linear model. This process was iterated backward in time for up to 600 days, generating a time series of regression results. Then, for each regression, we extracted the slope of the relationship (similar to the Peak signal detection method), and the model R^2. These values are then smoothed over time using a Generalized Additive Model (GAM) implemented in the mgcv R package (Wood

10

2017, version 1.9-3). The smoothed functions help identify the calendar days that have the greatest influence on seed production. This critical period was determined as the consecutive days in which the slope coefficient exceeds either the lower or upper quantiles (2.5% and 97.5% thresholds) calculated from all daily coefficients at the same time as the R^2 values exceed the upper quantile (97.5%) (Thackeray *et al.*, 2016).

P-spline regression

P-spline signal regression for cue identification was originally introduced by Roberts (2008) and follows a similar principle to the climate sensitivity profile method. However, instead of a two-step process, P-spline regression combines smoothing and coefficient estimation into a single step. This method regresses all 600 days of temperature against the response variable simultaneously, generating partial coefficients that describe the relationship between daily temperature and seed production. These coefficients are smoothed by penalizing differences between consecutive days to prevent overfitting. To handle the inclusion of numerous explanatory variables, coming from many time lags, P-spline regression incorporates a data-reduction phase using B-splines, which create a series of polynomial curves joined at predefined knots. The number of knots must be specified and is limited to one less than the sample size (Roberts, 2012). By combining B-splines with a difference penalty, the model applies P-splines (penalized B-splines) to enforce smooth transitions between coefficients. The penalty level is optimized through cross-validation to achieve the best balance between flexibility and interoperability. We implemented this method by using the mgcv R package, following the setup described in Roberts (2008); Roberts *et al.* (2015).

Time series length and cue identification

We assessed how the length of the time series affects the window identification by focusing on the longest time series available (> 50 years of observations, N = 15 populations). We divided each time series into four subsets of increasing length: 5, 10, 15, 20, 25, and 30 years (i.e., the subset of 30 years contains all smaller subsets). For each data subset and time series, we applied the four weather cue identification methods and extracted the window identified. This entire procedure was repeated 50 times to account for variation due to random selection of year blocks (this step is most significant for 5-year-long subsets). To facilitate comparisons, we summarized windows opening and closing to the median and inter-quantile range (IQR)

across each subset of time series length and weather cue method.

Cross validation of window identification and model performance

We evaluated the predictive performance of identified weather cues selected with each method by performing block cross-validation (Roberts *et al.*, 2017). We restricted this analysis to time series beginning after 1980 and randomly selected five populations for each of the three seed collection methods (seed count, seed trap, and visual crop assessment), yielding 15 populations in total. For each population, we extracted a continuous 30-year period and divided it into five equal blocks. Three blocks were randomly selected for model training and cue identification, while the remaining two blocks were used for validation by predicting seed production. This approach allowed us to evaluate model performance and the robustness of the selected weather cue across data subsets and collection methods.

Model accuracy was assessed using the coefficient of determination (R^2), based on comparisons between predicted and observed seed production in the validation dataset. We also calculated the normalized Root Mean Square Error (rRMSE = $RMSE/mean(observation)$), which reflects the average prediction error. An rRMSE near 0 indicates high accuracy, whereas values above 1 suggest performance worse than random noise.

Simulation study

We conducted a simulation study to assess how well the focal weather cue detection methods could identify a predefined cue window under varying levels of signal strength, expressed as the R^2 of the relationship between the cue window and annual seed production. We simulated seed production datasets using a known weather cue window, with temperature as the predictor, based on parameter ranges derived from our empirical analysis. Empirical distributions of model parameters—intercepts (α), slopes (β), and residual standard deviations (σ)—were obtained from 200 fitted models (50 time series × 4 cue identification methods). These parameters represent the estimated relationships between seed production and mean temperature over identified climatic windows.

The simulation model followed a linear regression form:

$$\log(\text{seed})_s = \alpha_s + \beta_s \times \text{Temperature}_{w,s} + \epsilon_s, \quad \epsilon_s \sim \mathcal{N}(0, \sigma_s^2)$$

We generated 1,000 datasets, each representing a simulated population s, using temperature values drawn from a predefined 10 days window (w)—June 10 to June 20— of the seedfall year (T0). Temperature values were scaled, and seed production was log-transformed to match the preprocessing used in the empirical models.

To explore a gradient of signal strength between temperature and seed production, we manipulated the residual variance σ_s, while drawing α_s and β_s from uniform distributions bounded by the empirical parameter ranges:

$$\alpha_s \sim U(\alpha_{min}, \alpha_{max}); \quad \beta_s \sim U(\beta_{min}, \beta_{max}); \quad \sigma_s \sim U(\sigma_{min}, \sigma_{max}) \tag{1}$$

By varying σ_s, we simulated datasets spanning a wide range of explanatory power, from very weak ($R^2 \approx 0$) to very strong ($R^2 \approx 0.99$) signal. This allowed us to assess the performance of each cue identification method under differing levels of signal detectability.

Results

Weather cue windows identified with the focal methods

Assuming the benchmark opening date for the weather cue in beech—i.e., the summer solstice (21st June)—is accurate, all methods performed reasonably well in identifying the cue across the 50 time series included in the study. Across all methods, the median estimated window opening was day 490, corresponding to 30th June (Fig. 2). However, the spread around this estimate varied by method: it was similarly narrow for the climate sensitivity profile, sliding window, and P-spline regression (each with $\pm$ 20 days), and substantially wider for peak signal detection (> 100 days). Interestingly, for the climate sensitivity profile, sliding window, and P-spline regression, the deviation from the benchmark date was asymmetrical—fewer simulations indicated window openings before the solstice—aligning with theoretical expectations (Journé *et al.*, 2024) (Fig. 2). In contrast, an apparently poorer performance of peak signal detection resulted from identifying the cue window in winter or spring of the seedfall year in 11 time series (Fig. S4).
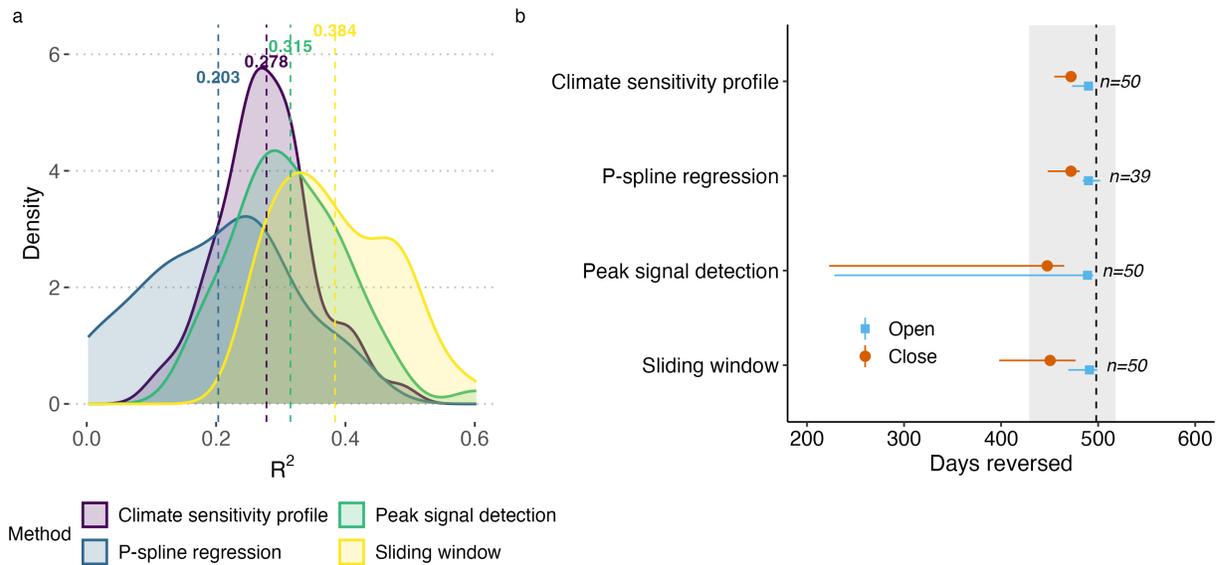
Figure 2: a) Distribution of coefficients of determination (R^2) across 50 populations of European beech for each weather cue identification method: climate sensitivity profile, P-spline regression, peak signal identification, and sliding window. Dashed lines indicate the average mean R^2 across populations for each method. b) Median window opening (blue) and closing (orange) dates for each method. Whiskers indicate the interquartile range. The black dashed line marks the summer solstice (21st June) in the year preceding seedfall (benchmark used in this study), while the grey shaded area highlights the summer months (June–August). N indicates the number of populations used per method; N was lower for P-spline regression due to time series that were too short or noisy to identify a reliable window.

The median window closing day estimated with the sliding window method was day 450 (9th August), closely matching that of peak signal detection (day 448, 11th August) (Fig. 2). In contrast, the climate sensitivity profile and P-spline regression yielded shorter windows, with median closing dates of day 472 (18th July) for both methods (Fig. 2). Variation around the median also differed among methods: it was narrowest for the climate sensitivity profile and P-spline regression (±15 days), broader for the sliding window method (±35 days), and widest for peak signal detection (±120 days) (Fig. 2). Peak signal detection showed the greatest deviation, with some runs producing windows that both opened and closed before the summer solstice (Fig. S4). Median opening and closing dates, along with their 95% interquartile ranges, are provided in Table S1. The best window identified for each population using the four methods is shown in Fig. S1, Fig. S2, Fig. S4, and Fig. S3. Note that, in contrast to the window-opening date, which appears anchored to the summer solstice, the closing date of the window is not associated with a known date (Journé *et al.*, 2024).

On average, the sliding window method provided the window with the highest model predictive performance (mean R^2 = 0.38), followed then by peak signal detection (mean R^2 = 0.32), climate sensitivity profile (mean R^2 = 0.28), and P-spline regression (mean R^2 = 0.20) (Fig. 2).

Time series length and cue identification

Reducing the sample size (i.e., shortening the time series) had a strong impact on the accuracy of the identified weather cue window, with methods differing in their sensitivity to data reduction. The climate sensitivity profile was the most robust, yielding median estimates for the window opening date that remained closely aligned with the summer solstice even when only 10 years of data were used (Figure 3). As expected, variation around the estimated dates was lowest when 25–30 years of data were included. For P-spline regression and the sliding window, at least 20 years of data were needed to achieve reasonably consistent estimates, while accurate alignment with the summer solstice was generally achieved with 30 years of data (Figure 3). In contrast, peak signal detection performed comparatively poorly across all sample sizes, including those with 25 or 30 years of data (Figure 3).

Model performance

Block cross-validation revealed that, on average, the climate sensitivity profile method achieved the highest predictive performance (mean R^2 = 0.18), followed by P-spline regression (R^2 = 0.17), the sliding window (R^2 = 0.12), and peak signal detection (R^2 = 0.11). Model performance varied significantly by seed collection method, particularly for the climate sensitivity profile and sliding window approaches, with seed traps consistently yielding higher accuracy (Figure 4, Figure S5). For the climate sensitivity profile, the mean R^2 was 0.22 when based on seed trap data, compared to 0.17 for seed counts and 0.13 for visual crop assessments. Similarly, for the sliding window method, seed traps produced a mean R^2 of 0.17, while seed counts and visual assessments yielded lower values (0.11 and 0.08, respectively).

Signal strength and cue detection

The simulation study showed that under very strong signal strength (R^2 > 0.75), both the sliding window and peak signal detection methods accurately recovered the predefined cue win-
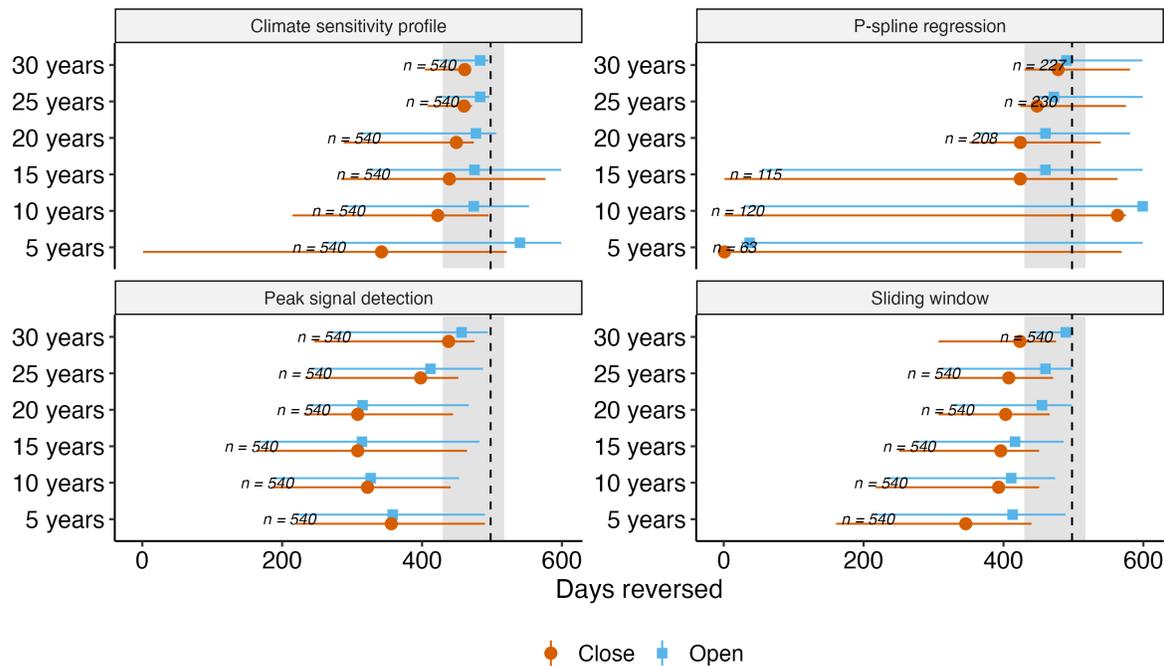
Figure 3: Effects of reducing time series length on the identified cue window. From the longest time series (>50 years of observation, N = 15), we randomly sampled 5, 10, 15, 20, 25, or 30 consecutive years, applied each method to identify the cue window, and repeated this process 50 times. Opening and closing dates identified in each iteration were aggregated to estimate medians and associated interquartile ranges (IQR) for each population.

dow, with median opening and closing dates closely matching the true values (Fig. 5). The climate sensitivity profile and p-spline regression also performed well in this scenario, although with greater variability around the estimates. At strong signal strength (R^2 = 0.5–0.75), the sliding window and peak signal detection methods remained robust, maintaining close alignment with the predefined window and showing only moderate increases in estimation error. Both methods continued to perform reasonably well under moderate signal strength (R^2 = 0.25–0.5), with median estimates still near the predefined dates and low to moderate error. When signal strength dropped below R^2 = 0.25, the accuracy of all methods declined, but the sliding window remained the most reliable, providing estimates still relatively close to the predefined window and with comparatively small errors. In contrast, the climate sensitivity profile was more sensitive to declining signal strength: it began to deviate from the true window already under strong signal conditions, showed increasing error under moderate signal strength, and failed to recover the correct window entirely under weak signals.
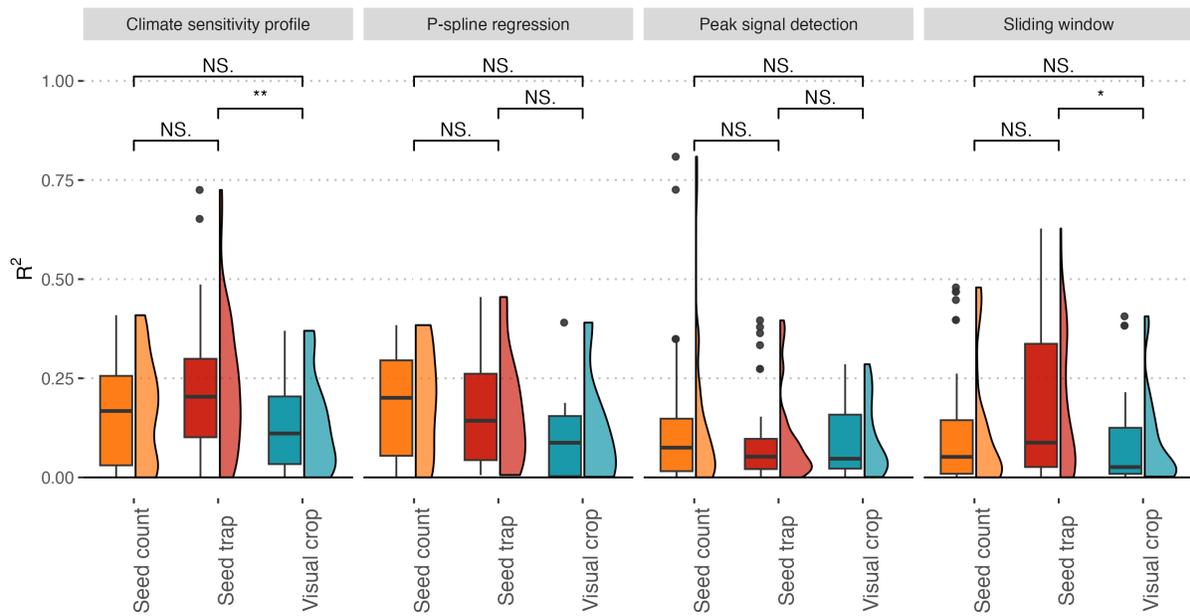
16

Figure 4: Model performance, measured by R^2, based on cross-validation using three blocks for training and two for validation, repeated over 10 iterations. Results are aggregated across a random sample of five time series per seed collection method (seed count, seed trap, and visual crop assessment). Pairwise differences between methods were tested using Wilcoxon tests, with significance levels denoted as "***" (p < 0.001), "**" (p < 0.01), "*" (p < 0.05), and "NS" for non-significant differences. Model accuracy based on nRMSE is shown on Figure S5.
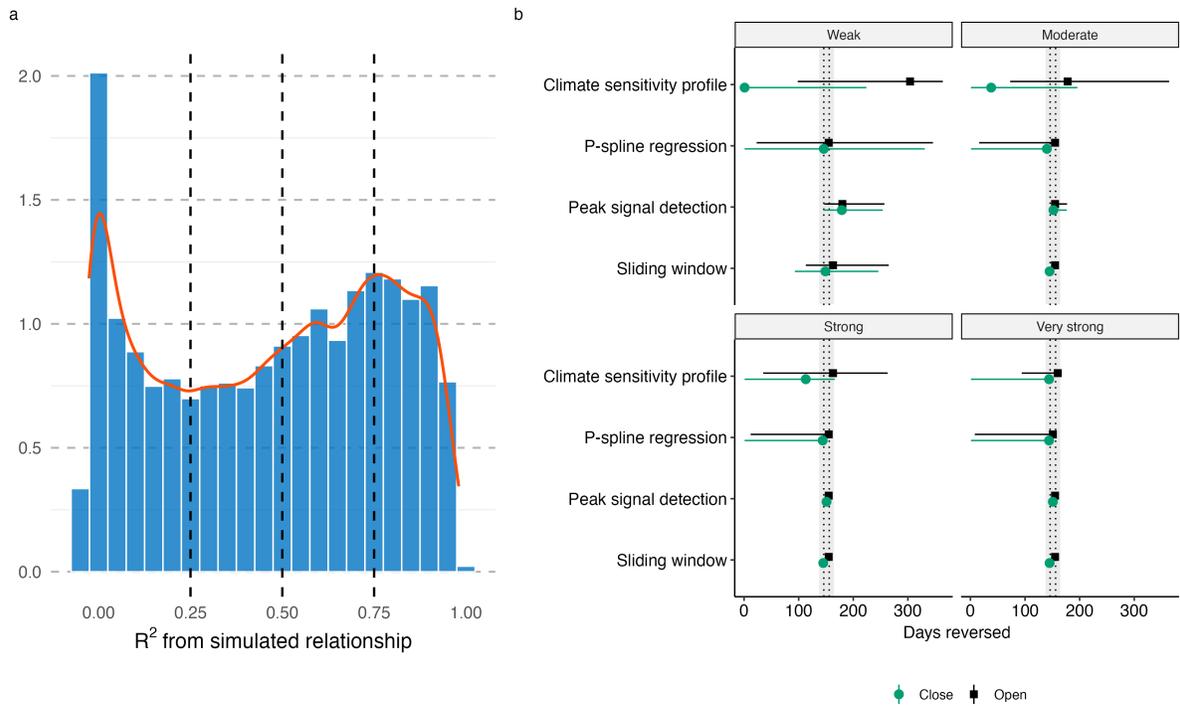
Figure 5: Simulation of window detection accuracy across the four window detection methods. We simulated 10,000 datasets, each spanning 40 years, by generating values from a linear model using randomly drawn parameters (intercept α, slope β, and error term σ) within ranges derived from empirical data. The predefined window influencing the biological response was fixed between days 150 and 160. (a) Distribution of simulated R^2 values, representing the strength of the relationship between the biological response and the weather cue. Simulations were categorized into four signal strength classes: weak ($R^2 < 0.25$), moderate (0.25–0.5), strong (0.5–0.75), and very strong ($R^2 > 0.75$). (b) Window detection performance across four methods, grouped by signal strength class. Points represent median estimated opening and closing dates; bars show the IQR (25th to 75th percentiles). The dashed lines indicate the predefined window range used in the simulations, with an additional 10-day margin around those dates highlighted with a shaded area.

Discussion

When applied to our dataset of 50 time series (each spanning at least 20 years, with a median length of 43 years), all four weather cue identification methods successfully detected the benchmark cue window, defined as the period just after the summer solstice (Journé *et al.*, 2024). This result, reinforced by our simulation study, demonstrates that these methods can reliably uncover biologically meaningful cues without requiring prior assumptions about their timing, an increasingly valuable capability in masting research. Hypersensitivity to weather cues is a central mechanism underlying mast seeding (Kelly, 1994; Bogdziewicz *et al.*, 2024b), and a substantial literature has examined correlations between weather variation and interannual seed production (Crone & Rapp, 2014). However, much of this work is constrained by the use of diverse *a priori* assumptions about which cues are relevant (Crone & Rapp, 2014). At the same time, recent findings also highlight within-species variation in the climatic drivers of seed production (Koenig *et al.*, 2020; Bogdziewicz *et al.*, 2023a; Fleurot *et al.*, 2023), which is increasingly recognized as important for explaining spatial synchrony in reproduction (Bogdziewicz *et al.*, 2023a), improving forecasting accuracy (Oberklamme *et al.*, 2025), and climate change biology (Bogdziewicz *et al.*, 2024b). Our results indicate that modern data-driven methods offer a framework for identifying key weather cues, providing an important step forward for both theoretical understanding and predictive modeling in masting systems.

Our comparison of cue identification methods highlights both their strengths and limitations, particularly in relation to data characteristics and signal clarity. A key limitation of peak signal detection appears to be its sensitivity to isolated, strong correlation peaks, even when these occur in biologically less likely periods. This tendency contributed to the larger deviations from the benchmark cue windows observed with peak signal detection in our analysis. In contrast, methods such as the sliding window and climate sensitivity profile are more robust to such anomalies. The sliding window approach systematically evaluates model fit across all possible time windows, while the climate sensitivity profile and P-spline regression smooth the signal using generalized additive models, reducing the influence of outliers.

Accurate detection of cue windows that would align with the benchmark strongly depends on data availability, with most methods requiring at least 20 years of observations to produce consistent results. This requirement is unsurprising given the long-term nature of masting, where years of high seed production are interspersed with multiple poor years, typ-

ically with an interval of 2–4 years between large-seeding years across species and populations (Qiu *et al.*, 2023; Kondrat *et al.*, 2025). Under such dynamics, even a 20-year time series may contain only a few large-seeding years, limiting the ability to robustly detect weather cues. The need for long-term data is well-established in the field. For instance, widely used metrics to measure interannual variation in seed production, such as the coefficient of variation (CV), require 10–20 years of data before estimates stabilize (Lobry *et al.*, 2023; Foest *et al.*, 2025b). In this context, our findings offer new guidance by showing that even 15 years of data may be insufficient for cue identification, particularly when the underlying signal is weak, stressing the importance of long-term monitoring for understanding masting dynamics. Gladly, database compilations are increasingly available (Clark *et al.*, 2021; Hacket-Pain *et al.*, 2022; Nigro *et al.*, 2024), making application of such restrictions in analysis (e.g., 15+ years of data) possible.

Block cross-validation showed that cue windows identified using the climate sensitivity profile and P-spline regression consistently yielded the highest predictive accuracy, particularly when models were trained on seed trap data. Accurately determining the timing of weather cues is a critical step in improving masting forecasts (Journé *et al.*, 2023; Fleurot *et al.*, 2023; Wion *et al.*, 2025; Oberklamme *et al.*, 2025). Forecast performance improves when weather predictors are drawn from biologically relevant periods and paired with reliable data on past seed production (Journé *et al.*, 2023; Oberklamme *et al.*, 2025). Among the methods tested, climate sensitivity profile and P-spline regression produced the most informative cues, likely due to their capacity to smooth short-term variability and isolate consistent weather drivers. Additionally, seed–trap–based monitoring, previously shown to reduce uncertainty in reproductive estimates used for forecasting (Journé *et al.*, 2023), was associated with stronger model performance in our analysis. Together, these results highlight cue identification as a foundation for robust mast forecasts. Nonetheless, for practical reasons (i.e., lack of a clear benchmark in other species), our conclusions are based on a single species. Thus, our results, including the superiority of seed trap monitoring, may not generalize across taxa or ecological contexts, highlighting the need for broader testing across systems.

In summary, all four cue window detection methods performed well, but with important variation in performance across tests (Table 2). In terms of identifying the benchmark across all data from all 50 populations studied, peak signal detection performed less reliably, identifying windows far from the benchmark in several time series. The sample size

reduction experiment showed that 20 years of data are recommended for consistent and accurate cue identification. While acceptable deviations will vary depending on study goals, researchers working with only 10–15 years of data may consider using the climate sensitivity profile method, which remained relatively robust under such constraints. Nevertheless, this approach should be applied cautiously when the relationship between weather and seed production is weak, as the climate sensitivity profile may then fail to recover the correct window. Taken together, our findings suggest that, depending on the study goals, the climate sensitivity profile or sliding window methods offer the best balance of accuracy and robustness (Table 2). The sliding window approach was more tolerant of reduced data availability but tended to produce cues with lower predictive power. P-spline regression also performed well under favorable conditions but showed reduced reliability when the sample size was limited. We provide an R compendium to facilitate applications of these methods in mast seeding research.

Table 2: Summary of the cue window detection methods across various tests performed in this study.

Method	General accuracy vs benchmark	Performance under limited time series length	Predictive power of identified cues (Mean R^2)	Detecting a predefined cue window under varying signal strength
Sliding window	Very good	Good	0.12	Very good
Peak signal detection	Failed in several runs	Weak	0.11	Very good
Climate sensitivity profile	Very good	Very good	0.18	Moderate
P-spline regression	Very good	Moderate	0.17	Moderate

References

Abe, T., Tachiki, Y., Kon, H., Nagasaka, A., Onodera, K., Minamino, K. *et al.* (2016). Parameterisation and validation of a resource budget model for masting using spatiotemporal flowering data of individual trees. *Ecology Letters*, 19, 1129–1139.

Ascoli, D., Hacket-Pain, A., LaMontagne, J.M., Cardil, A., Conedera, M., Maringer, J. *et al.* (2020). Climate teleconnections synchronize picea glauca masting and fire disturbance: Evidence for a fire-related form of environmental prediction. *Journal of Ecology*, 108, 1186–1198.

Bailey, L.D. & van de Pol, M. (2016). Climwin: An r toolbox for climate window analysis. *PLoS ONE*, 11.

Bajocco, S., Ferrara, C., Bascietto, M., Alivernini, A., Chirichella, R., Cutini, A. *et al.* (2021). Characterizing the climatic niche of mast seeding in beech: Evidences of trade-offs between vegetation growth and seed production. *Ecological Indicators*, 121.

Bogdziewicz, M., Ascoli, D., Hacket-Pain, A., Koenig, W.D., Pearse, I., Pesendorfer, M. *et al.* (2020a). From theory to experiments for testing the proximate mechanisms of mast seeding: an agenda for an experimental ecology. *Ecology Letters*, 23, 210–220.

Bogdziewicz, M., Chybicki, I., Szymkowiak, J., Ulaszewski, B., Burczyk, J., Szarek-Łukaszewska, G. *et al.* (2024a). Masting and efficient production of seedlings: Balancing costs of variation through synchronised fruiting. *Ecology letters*, 27.

Bogdziewicz, M., Hacket-Pain, A., Kelly, D., Thomas, P.A., Lageard, J. & Tanentzap, A.J. (2021). Climate warming causes mast seeding to break down by reducing sensitivity to weather cues. *Glob Change Biol*, 27, 1952–1961.

Bogdziewicz, M., Journé, V., Hacket-Pain, A. & Szymkowiak, J. (2023a). Mechanisms driving interspecific variation in regional synchrony of trees reproduction. *Ecol Lett*, 26, 754–764.

Bogdziewicz, M., Kelly, D., Ascoli, D., Caignard, T., Chianucci, F., Crone, E.E. *et al.* (2024b). Evolutionary ecology of masting: mechanisms, models, and climate change. *Trends in Ecology and Evolution*, 39, 851–862.

Bogdziewicz, M., Kelly, D., Tanentzap, A.J., Thomas, P., Foest, J., Lageard, J. *et al.* (2023b). Reproductive collapse in european beech results from declining pollination efficiency in large trees. *Glob Change Biol*, 29, 4595–4604.

Bogdziewicz, M., Kelly, D., Thomas, P.A., Lageard, J.G.A. & Hacket-Pain, A. (2020b). Climate warming disrupts mast seeding and its fitness benefits in european beech. *Nat Plants*, 6, 88–94.

Bogdziewicz, M., Kelly, D., Zwolak, R., Szymkowiak, J. & Hacket-Pain, A. (2025). Dynamics, mechanisms, and consequences of mast seeding. *Annual Reviews in Ecology, Evolution, and Systematics*.

Bogdziewicz, M., Szymkowiak, J., Fernández-Martínez, M., Peñuelas, J. & Espelta, J.M. (2019). The effects of local climate on the correlation between weather and seed production differ in two species with contrasting masting habit. *Agricultural and Forest Meteorology*, 268, 109–115.

Brakel, J.v. (2014). Robust peak detection algorithm using z-scores. https://stackoverflow.com/questions/22583391/peak-signal-detection-in-realtime-timeseries-data/2264036222640362.

Clark, J.S., Andrus, R., Aubry-Kientz, M., Bergeron, Y., Bogdziewicz, M., Bragg, D.C. *et al.* (2021). Continent-wide tree fecundity driven by indirect climate effects. *Nature Communications*, 12, 1242.

Clark, J.S., Nuñez, C.L. & Tomasek, B. (2019). Foodwebs based on unreliable foundations: spatiotemporal masting merged with consumer movement, storage, and diet. *Ecological Monographs*, 89.

Cornes, R.C., van der Schrier, G., van den Besselaar, E.J. & Jones, P.D. (2018). An ensemble version of the e-obs temperature and precipitation data sets. *Journal of Geophysical Research: Atmospheres*, 123, 9391–9409.

Crone, E.E. & Rapp, J.M. (2014). Resource depletion, pollen coupling, and the ecology of mast seeding. *Annals of the New York Academy of Sciences*, 1322, 21–34.

Fleurot, E., Keurinck, L., Boulanger, V., Debias, F., Delpierre, N., Delzon, S. *et al.* (2024). Reconciling Pollen Limitation Theories: Insights From Temperate Oak Masting. *Ecology Letters*, 27.

Fleurot, E., Lobry, J.R., Boulanger, V., Debias, F., Mermet-Bouvier, C., Caignard, T. *et al.* (2023). Oak masting drivers vary between populations depending on their climatic environments. *Curr Biol*, 33, 1117–1124.

Foest, J.J., Bogdziewicz, M., Caignard, T., Hadad, M., Thomas, P.A. & Hacket-Pain, A. (2025a). Comparing two ground-based seed count methods and their effect on masting metrics. *Forest Ecology and Management*, 581, 122551.

Foest, J.J., Bogdziewicz, M., Pesendorfer, M.B., Ascoli, D., Cutini, A., Nussbaumer, A. *et al.* (2024). Widespread breakdown in masting in european beech due to rising summer temperatures. *Glob Change Biol*, 30, e17307.

Foest, J.J., Caignard, T., Pearse, I.S., Bogdziewicz, M. & Hacket-Pain, A. (2025b). Intraspecific variation in masting across climate gradients is inconsistent with the environmental stress hypothesis. *Ecology*, 106.

Hacket-Pain, A., Foest, J.J., Pearse, I.S., LaMontagne, J.M., Koenig, W.D., Vacchiano, G. *et al.* (2022). Mastree+: Time-series of plant reproductive effort from six continents. *Glob Change Biol*, 28, 3066–3082.

Hacket-Pain, A., Szymkowiak, J., Journé, V., Barczyk, M.K., Thomas, P.A., Lageard, J.G.A. *et al.* (2025). Growth decline in european beech associated with temperature-driven increase in reproductive allocation. *Proceedings of the National Academy of Sciences*, 122.

Hacket-Pain, A.J., Ascoli, D., Vacchiano, G., Biondi, F., Cavin, L., Conedera, M. *et al.* (2018). Climatically controlled reproduction drives interannual growth variability in a temperate tree species. *Ecol Lett*, 21, 1833–1844.

Janzen, D.H. (1971). Seed predation by animals. *Annual Review of Ecology and Systematics*, 220, 465–492.

Journé, V., Hacket-Pain, A., Oberklammer, I., Pesendorfer, M.B. & Bogdziewicz, M. (2023). Forecasting seed production in perennial plants: identifying challenges and charting a path forward. *New Phytol*, 239, 466–476.

Journé, V., Szymkowiak, J., Foest, J., Hacket-Pain, A., Kelly, D. & Bogdziewicz, M. (2024). Summer solstice orchestrates the subcontinental-scale synchrony of mast seeding. *Nat Plants*, 10, 367–373.

Kelly, D. (1994). The evolutionary ecology of mast seeding. *Trends Ecol Evol*, 9, 465–470.

Kelly, D., Geldenhuis, A., James, A., Holland, E.P., Plank, M.J., Brockie, R.E. *et al.* (2013). Of mast and mean: differential-temperature cue makes mast seeding insensitive to climate change. *Ecol Lett*, 16, 90–98.

Kelly, D., Hart, D.E. & Allen, R.B. (2001). Evaluating the wind pollination benefits of mast seeding. *Ecology*, 82, 117–126.

Koenig, W.D. (2021). A brief history of masting research. *Philos Trans R Soc Lond B Biol Sci*, 376.

Koenig, W.D., Knops, J.M. & Carmen, W.J. (2020). Intraspecific variation in the relationship between weather and masting behavior in valley oak, quercus lobata. *Canadian Journal of Forest Research*, 50.

Koenig, W.D., Knops, J.M., Carmen, W.J. & Pearse, I.S. (2015). What drives masting? the phenological synchrony hypothesis. *Ecology*, 96, 184–192.

Kondrat, K., Szymkowiak, J., Hacket-Pain, A., Shibata, M., Saitoh, T., Foest, J. *et al.* (2025). Short reproductive periods dominate mast seeding across diverse tree species. *EcoEvoRxiv*.

LaMontagne, J.M., Redmond, M.D., Wion, A.P. & Greene, D.F. (2021). An assessment of temporal variability in mast seeding of north american pinaceae. *Philos Trans R Soc Lond B Biol Sci*, 376.

Lee, H., Otero-Leon, D., Dong, H., Stringfellow, E.J. & Jalali, M.S. (2024). Uncovering patterns in overdose deaths: An analysis of spike identification in fatal drug overdose data in massachusetts, 2017-2023. *Public Health Reports*.

Lobry, J.R., Bel-Venner, M.C., Bogdziewicz, M., Hacket-Pain, A. & Venner, S. (2023). The cv is dead, long live the cv! *Methods Ecol Evol*, 14, 2780–2786.

Maag, N., Korner-Nievergelt, F., Szymkowiak, J., Hałas, N., Maziarz, M., Neubauer, G. *et al.* (2024). Wood warbler population dynamics in response to mast seeding regimes in europe. *Ecology*, 105.

Michaud, T.J., Pearse, I.S., Kauserud, H., Andrew, C.J. & Kennedy, P.G. (2024). Mast seeding in european beech (fagus sylvatica l.) is associated with reduced fungal sporocarp production and community diversity. *Ecology Letters*, 27.

Nigro, K.M., Barton, J.H., Macias, D., Chaudhary, V.B., Pearse, I.S., Bell, D.M. *et al.* (2024). Co-mast: Harmonized seed production data for woody plants across us long-term research sites. *Ecology*.

Numata, S., Yamaguchi, K., Shimizu, M., Sakurai, G., Morimoto, A., Alias, N. *et al.* (2022). Impacts of climate change on reproductive phenology in tropical rainforests of southeast asia. *Comm Biol*, 5, 311.

Nussbaumer, A., Waldner, P., Apuhtin, V., Aytar, F., Benham, S., Bussotti, F. *et al.* (2018). Impact of weather cues and resource dynamics on mast occurrence in the main forest tree species in europe. *Forest Ecology and Management*, 429, 336–350.

Oberklamme, I., Gratzer, G., Schueler, S., Konrad, Heino, Hacket-Pain *et al.* (2025). See(d)ing the seeds-toward weather-based forecasting of annual seed production in six european forest tree species. *EcoEvoRxiv*.

Pearse, I.S., Koenig, W.D. & Kelly, D. (2016). Mechanisms of mast seeding: resources, weather, cues, and selection. *New Phytol*, 212, 546–562.

Pearse, I.S., Koenig, W.D. & Knops, J.M.H. (2014). Cues versus proximate drivers: testing the mechanism behind masting behavior. *Source: Oikos*, 123, 179–184.

Pearse, I.S., Wion, A.P., Gonzalez, A.D. & Pesendorfer, M.B. (2021). Understanding mast seeding for conservation and land management. *Philos Trans R Soc Lond B Biol Sci*, 376.

Pesendorfer, M.B., Ascoli, D., Bogdziewicz, M., Hacket-Pain, A., Pearse, I.S. & Vacchiano, G. (2021). The ecology and evolution of synchronized reproduction in long-lived plants. *Philos Trans R Soc Lond B Biol Sci*, 376, 20200369.

Pesendorfer, M.B., Bogdziewicz, M., Szymkowiak, J., Borowski, Z., Kantorowicz, W., Espelta, J.M. *et al.* (2020). Investigating the relationship between climate, stand age, and temporal trends in masting behavior of european forest trees. *Global Change Biology*, 26, 1654–1667.

Piovesan, G. & Adams, J.M. (2011). Masting behaviour in beech: linking reproduction and climatic variation. *https://doi.org/10.1139/b01-089*, 79, 1039–1047.

van de Pol, M., Bailey, L.D., McLean, N., Rijsdijk, L., Lawson, C.R. & Brouwer, L. (2016). Identifying the best climatic predictors in ecology and evolution. *Methods in Ecology and Evolution*, 7, 1246–1257.

Pérez-Ramos, I.M., Ourcival, J.M., Limousin, J.M. & Rambal, S. (2010). Mast seeding under increasing drought: results from a long-term data set and from a rainfall exclusion experiment. *Ecology*, 91, 3057–3068.

Qiu, T., Aravena, M.C., Ascoli, D., Bergeron, Y., Bogdziewicz, M., Boivin, T. *et al.* (2023). Masting is uncommon in trees that depend on mutualist dispersers in the context of global climate and fertility gradients. *Nat Plants*, 9, 1044–1056.

Redmond, M.D., Davis, T.S., Ferrenberg, S. & Wion, A.P. (2019). Resource allocation trade-offs in a mast-seeding conifer: Piñon pine prioritizes reproduction over defence. *AoB PLANTS*, 11.

Roberts, A.M. (2008). Exploring relationships between phenological and weather data using smoothing. *International Journal of Biometeorology*, 52, 463–470.

Roberts, A.M., Tansey, C., Smithers, R.J. & Phillimore, A.B. (2015). Predicting a change in the order of spring phenology in temperate forests. *Global Change Biology*, 21, 2603–2611.

Roberts, A.M.I. (2012). Comparison of regression methods for phenology. *International Journal of Biometeorology*, 56, 707–717.

Roberts, D.R., Bahn, V., Ciuti, S., Boyce, M.S., Elith, J., Guillera-Arroita, G. *et al.* (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40, 913–929.

Samarth, Lee, R., Kelly, D., Turnbull, M.H., Macknight, R.C., Poole, A.M. *et al.* (2021). Molecular control of the floral transition in the mast seeding plant celmisia lyallii (asteraceae). *Mol Ecol*, 30, 1846–1863.

Schermer, E., Bel-Venner, M.C., Gaillard, J.M., Dray, S., Boulanger, V., Roncé, I.L. *et al.* (2020). Flower phenology as a disruptor of the fruiting dynamics in temperate oak species. *New Phytol*, 225, 1181–1192.

Shibata, M., Masaki, T., Yagihashi, T., Shimada, T. & Saitoh, T. (2020). Decadal changes in masting behaviour of oak trees with rising temperature. *J Ecol*, 108, 1088–1100.

Simmonds, E.G., Cole, E.F. & Sheldon, B.C. (2019). Cue identification in phenology: A case study of the predictive performance of current statistical tools. *Journal of Animal Ecology*, 88, 1428–1440.

Szymkowiak, J., Foest, J., Hacket-Pain, A., Journé, V., Ascoli, D. & Bogdziewicz, M. (2024). Tail-dependence of masting synchrony results in continent-wide seed scarcity. *Ecology Letters*, 27.

Thackeray, S.J., Henrys, P.A., Hemming, D., Bell, J.R., Botham, M.S., Burthe, S. *et al.* (2016). Phenological sensitivity to climate across taxa and trophic levels. *Nature*, 535, 241–245.

Vacchiano, G., Hacket-Pain, A., Turco, M., Motta, R., Maringer, J., Conedera, M. *et al.* (2017). Spatial patterns and broad-scale weather cues of beech mast seeding in europe. *New Phytologist*, 215, 595–608.

Widick, I.V., Strong, C., LaMontagne, J.M., Young, M.A. & Zuckerberg, B. (2025). Continent-wide patterns of climate and mast seeding entrain boreal bird irruptions. *Global Change Biology*, 31.

Wion, A.P., Pearse, I.S., Broxson, M. & Redmond, M.D. (2025). Mast hindcasts reveal pervasive effects of extreme drought on a foundational conifer species. *New Phytologist*, 246, 450–460.

Wood, S.N. (2017). *Generalized Additive Models: An Introduction with R*. 2nd edn. Chapman and Hall/CRC.

Yukich-Clendon, O.M.M., Carpenter, J.K., Kelly, D., Timoti, P., Burns, B.R., Boswijk, G. *et al.* (2023). Global change explains reduced seeding in a widespread new zealand tree: indigenous tūhoe knowledge informs mechanistic analysis. *Front For Glob Change*, 6.

Zwolak, R., Celebias, P. & Bogdziewicz, M. (2022). Global patterns in the predator satiation effect of masting: A meta-analysis. *PNAS*, 119.

Acknowledgements

Competing interests

The authors declare no competing interests

Supplementary Materials

Table S1 – S6

Fig S1 - S6

Supplementary material

Comparing statistical methods for detecting climatic drivers of mast seeding, Journé et al.

Supplementary Tables and Figures

Table S1: Median and interquartile range (25th and 75th percentiles) of window opening and closing dates identified by each weather cue detection method (n = 50 sites per method, except for P-spline regression with n = 39 due to convergence issues). Model performance (R^2) is also reported as the median and interquartile range across all time series.

Method	Window (Open-Close)	Median [IQR]	R^2
Climate sensitivity profile	Opening	490 [473, 492]	0.27 [0.23, 0.32]
	Closing	472 [455, 478]	
P-spline regression	Opening	490 [484, 502]	0.23 [0.12, 0.28]
	Closing	472 [448, 481]	
Peak signal detection	Opening	489 [228, 495]	0.31 [0.26, 0.37]
	Closing	448 [223, 465]	
Sliding window	Opening	491 [469, 499]	0.37 [0.31, 0.46]
	Closing	450 [398, 477]	

Table S2: Median and interquartile range (25th and 75th percentiles) of window opening and closing dates identified by each weather cue detection method, shown in relation to time series length.

Method	Length Time series	Window Open	Window Close
Climate sensitivity profile	5	540 [249, 599]	341 [1, 520]
	10	474 [290, 553]	422 [215, 495]
	15	475 [326, 599]	439 [284, 576]
	20	477 [312, 506]	449 [288, 474]
	25	483 [424, 496]	460 [408, 471]
	30	483 [420, 495]	461 [404, 470]
P-spline regression	5	37 [25, 599]	1 [1, 569]
	10	599 [30, 599]	563 [1, 575]
	15	460 [55, 599]	424 [1, 563]
	20	460 [363, 581]	424 [351, 539]
	25	472 [448, 599]	448 [424, 575]
	30	490 [460, 599]	478 [430, 581]
Peak signal detection	5	358 [220, 490]	356 [220, 490]
	10	326 [188, 453]	322 [184, 441]
	15	314 [165, 482]	308 [163, 464]
	20	315 [240, 467]	308 [231, 444]
	25	412 [242, 487]	398 [233, 452]
	30	456 [269, 494]	438 [246, 475]
Sliding window	5	413 [216, 489]	346 [161, 440]
	10	411 [234, 474]	393 [217, 451]
	15	416 [275, 486]	396 [250, 451]
	20	455 [330, 497]	403 [307, 466]
	25	460 [310, 498]	408 [302, 471]
	30	489 [442, 500]	424 [307, 475]

Table S3: Summary of block cross-validation results showing window detection accuracy and model performance across weather cue detection methods and seed collection techniques.

| Method	Seed collection	Window	Median [IQR]	R^2	$nRMSE$
Climate sensitivity profile	Seed count	Closing	476 [453, 494]	0.17 ± 0.13	1.05 ± 0.21
		Opening	501 [485, 526]		
	Seed trap	Closing	468 [464, 475]	0.22 ± 0.17	0.99 ± 0.21
		Opening	497 [491, 499]		
	Visual crop	Closing	465 [426, 471]	0.13 ± 0.12	1.02 ± 0.23
		Opening	495 [486, 498]		
P-spline regression	Seed count	Closing	563 [96, 569]	0.19 ± 0.14	1.08 ± 0.11
		Opening	599 [116, 599]		
	Seed trap	Closing	394 [1, 558]	0.17 ± 0.14	1.01 ± 0.12
		Opening	406 [24, 599]		
	Visual crop	Closing	545 [1, 560]	0.11 ± 0.14	1.02 ± 0.06
		Opening	599 [31, 599]		
Peak signal detection	Seed count	Closing	256 [209, 420]	0.13 ± 0.18	1.18 ± 0.29
		Opening	265 [214, 427]		
	Seed trap	Closing	280 [197, 442]	0.09 ± 0.11	1.12 ± 0.21
		Opening	283 [202, 462]		
	Visual crop	Closing	359 [188, 450]	0.1 ± 0.09	1.17 ± 0.28
		Opening	359 [191, 494]		
Sliding window	Seed count	Closing	260 [174, 482]	0.11 ± 0.15	1.29 ± 0.3
		Opening	270 [193, 486]		
	Seed trap	Closing	462 [217, 479]	0.17 ± 0.17	1.15 ± 0.3
		Opening	487 [230, 510]		
	Visual crop	Closing	417 [300, 494]	0.08 ± 0.11	1.36 ± 0.35
		Opening	433 [392, 497]		
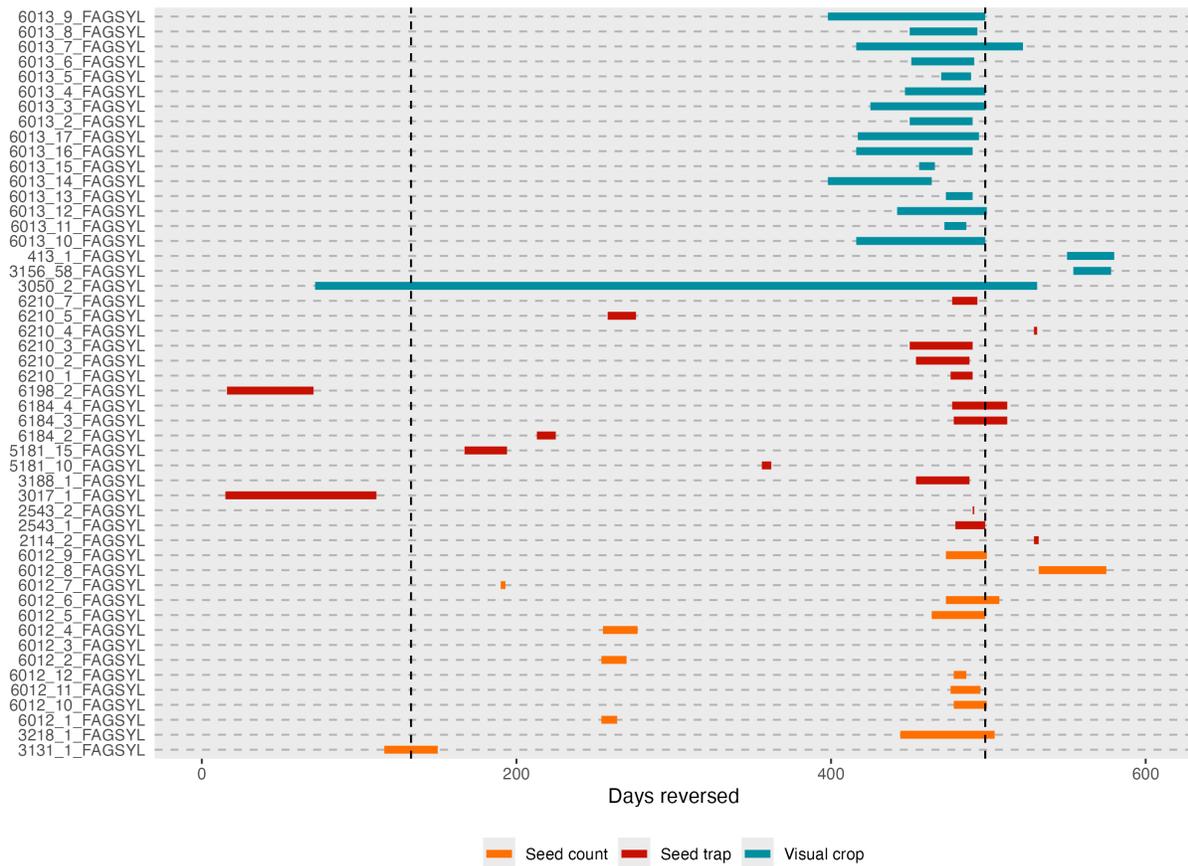
Figure S1: Identified cue windows for 50 European beech populations based on the sliding window method. The black dashed lines indicate the summer solstice of the seedfall year (133 days before the reference date, 1st November) and of the previous year (498 days before). Colors represent the seed collection method used: seed count (orange), seed trap (red), and visual crop assessment (blue).
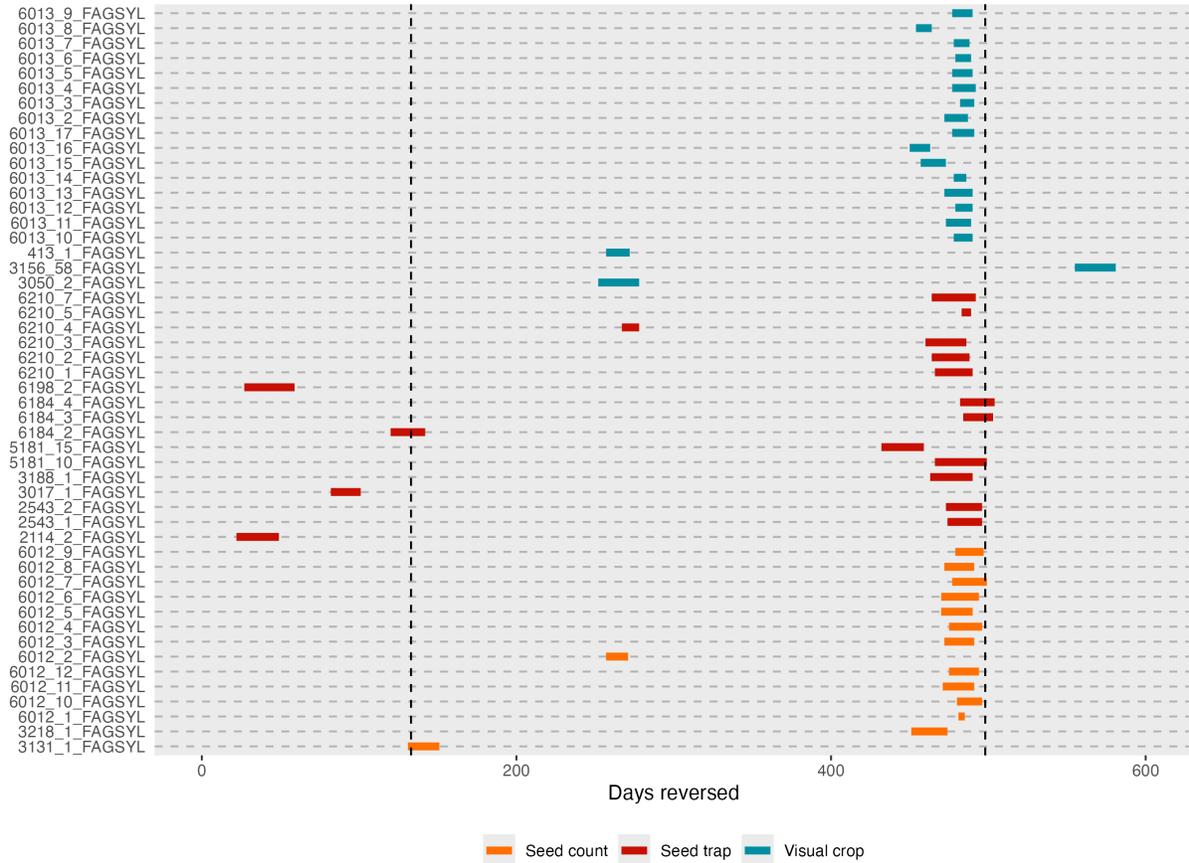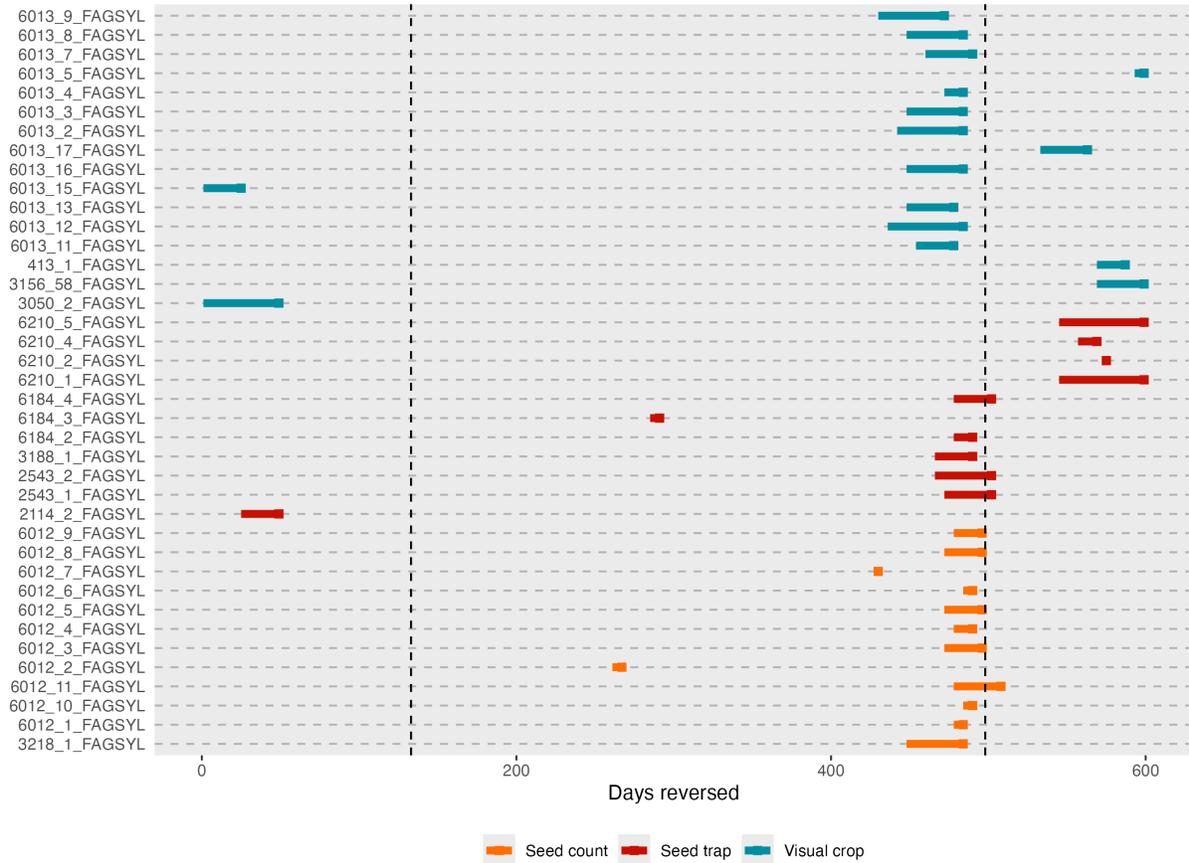
Figure S2: Identified cue windows for 50 European beech populations based on the climate sensitivity profile. The black dashed lines indicate the summer solstice of the seedfall year (133 days before the reference date, 1st November) and of the previous year (498 days before). Colors represent the seed collection method used: seed count (orange), seed trap (red), and visual crop assessment (blue).

Figure S3: Identified cue windows for 44 European beech populations based on the P-spline regression method. PSR failed to detect a window for 6 additional sites. The black dashed lines indicate the summer solstice of the seedfall year (133 days before the reference date, 1st November) and of the previous year (498 days before). Colors represent the seed collection method used: seed count (orange), seed trap (red), and visual crop assessment (blue).
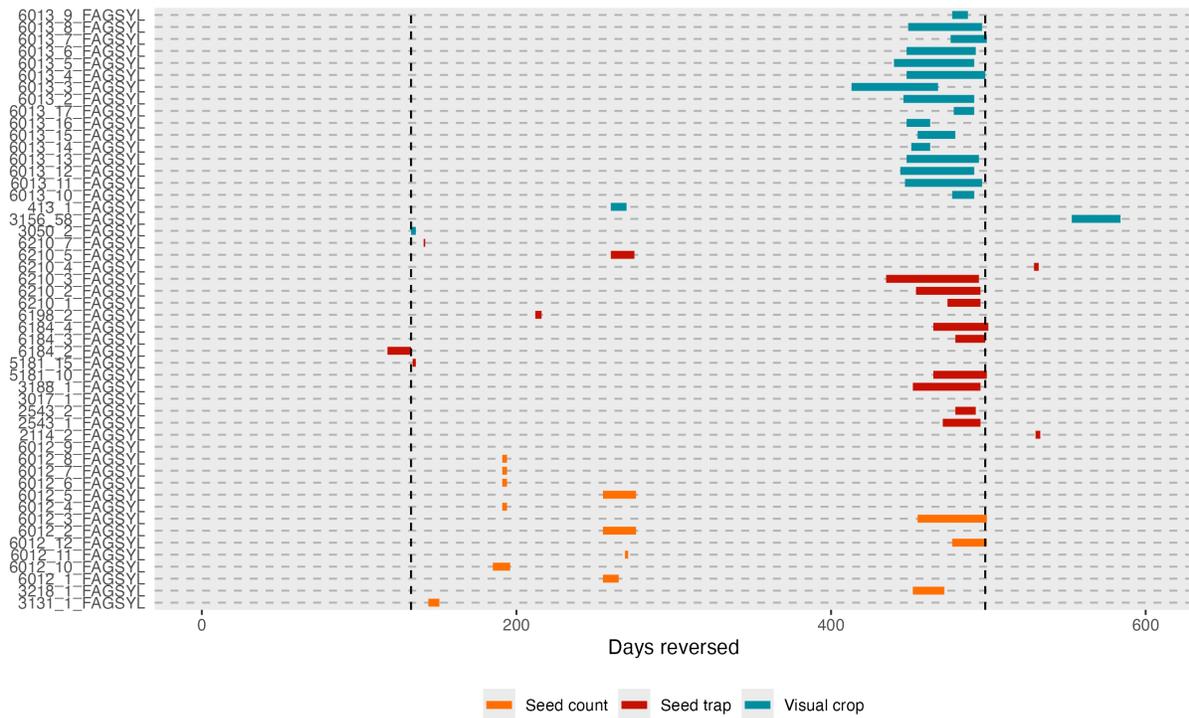
Figure S4: Identified cue windows for 50 European beech populations based on the peak signal detection method. The black dashed lines indicate the summer solstice of the seedfall year (133 days before the reference date, 1st November) and of the previous year (498 days before). Colors represent the seed collection method used: seed count (orange), seed trap (red), and visual crop assessment (blue).
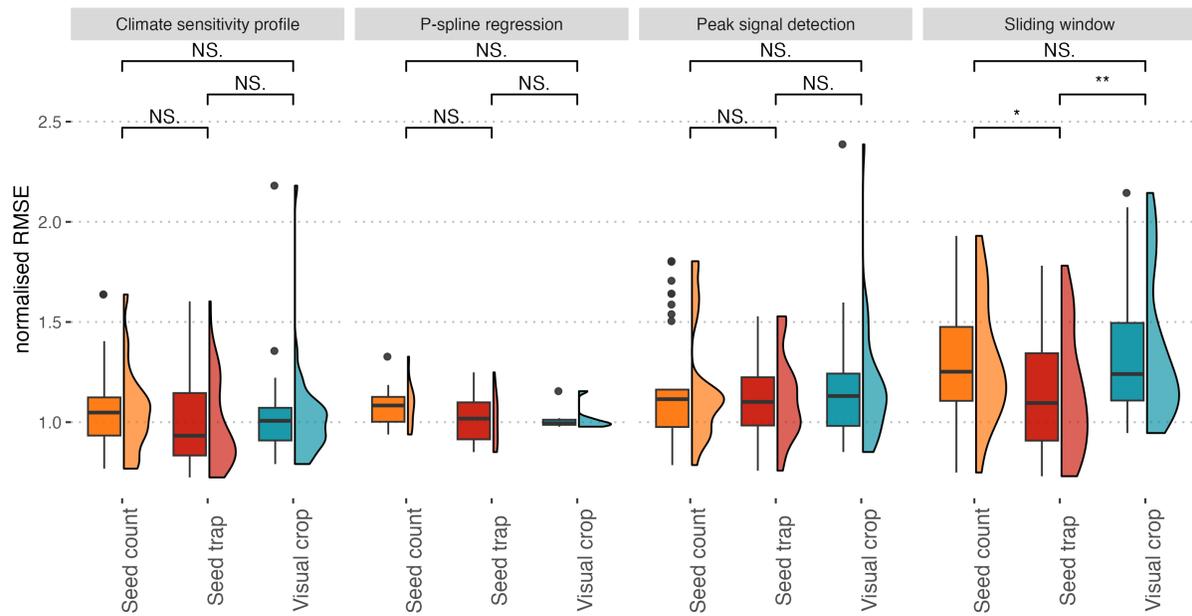
Figure S5: Model performance, measured normalised RMSE, based on cross-validation using three blocks for training and two for validation, repeated over 10 iterations. Results are aggregated across a random sample of five time series per seed collection method (seed count, seed trap, and visual crop assessment). Pairwise differences between methods were tested using Wilcoxon tests, with significance levels denoted as "***" (p < 0.001), "**" (p < 0.01), "*" (p < 0.05), and "NS" for non-significant differences.
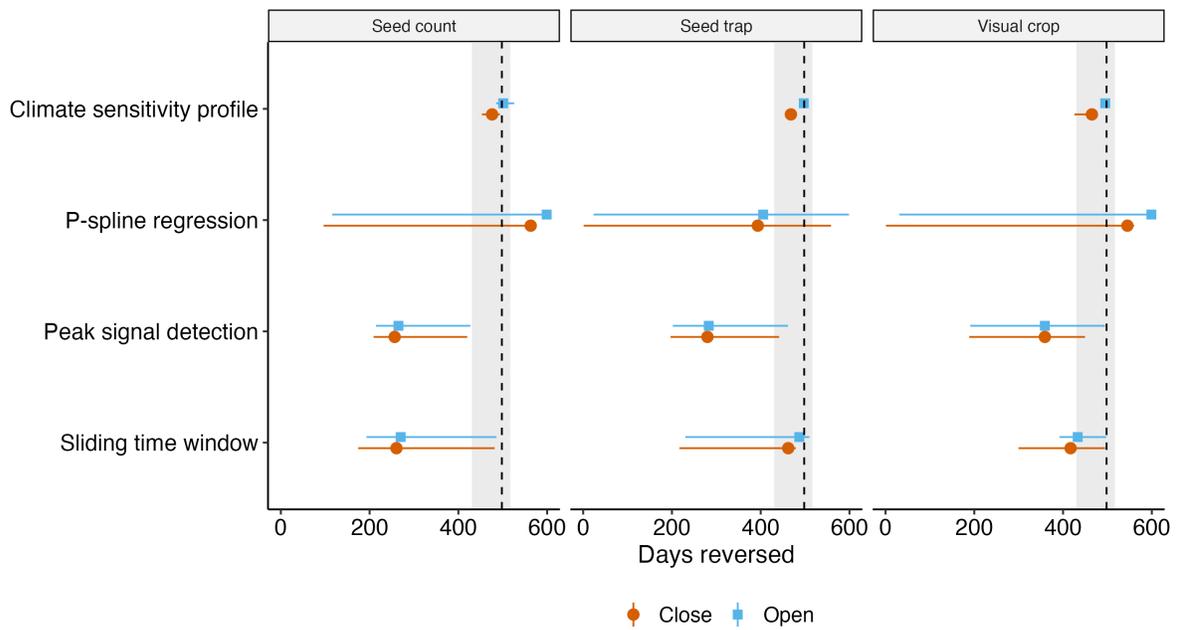
Figure S6: Robustness of window opening and closing identification using block cross validation. We aggregated opening and closing window, iterated over 10 times for each collection methods. We reported median and IQR for both window opening and closing. Grey shaded area corresponds to June-July of the previous year with vertical dotted line for summer solstice.