Inferring reproductive phenology and success from proportions of juveniles in population monitoring.

Paul Cuchot[1*], Luis-Miguel Chevin[1]

[1]CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

*Corresponding author: paul.cuchot@gmail.com

Abstract

1. Phenological shifts caused by climate change are increasingly documented in wild populations by the widespread collection of datasets on reproductive timing and success. These phenological events are often inferred by examining changes in population abundance and age structure throughout the breeding season. However, the quantitative relationship between the observed proportion of juveniles over time and the underlying distribution of breeding times (e.g., onset of reproduction) and average reproductive success is often not explicitly addressed. In addition, potential biases introduced by selection on reproductive phenology or by the sampling design can affect our inference of reproductive phenology and success.
2. In this study, taking the example of bird monitoring, we proposed an analytical model to relate the proportion of juveniles in counts (e.g., mist-net captures) to the distribution of fledging dates and mean reproductive success in the population. We then infer laying dates from fledging dates, accounting for putative selection through fertility and/or juvenile survival to fledging. We simulated varying levels of variance, selection strength, and sampling effort.
3. Our simulation results show that our approach is able to recover the true mean and variance of laying dates and the mean reproductive success under ideal conditions

(large sampling effort, no selection). It notably corrects for the fact that the mean fledging time lags behind the inflection point in the proportion of sampled juveniles, all the more so as laying date variance and reproductive success are high. Selection for earlier breeding systematically biases the estimates of mean laying dates, but we show how this bias can be corrected if information on selection strength is available. Multi-site analyses reveal that low sampling effort and high within-site variation can introduce further biases, but also suggest that four sampling sessions with reasonable effort per year provides reasonable estimates.

4. These findings underscore the importance of explicitly modeling the population processes (including possibly selection), and carefully planning sampling designs, to improve the accuracy of phenological estimates from population monitorings, and better interpret climate-driven changes in wild populations.

Key words: Phenological shift, breeding, fledging, sampling

Introduction

Phenological shifts represent a major biological response to climate change (Parmesan & Yohe, 2003). Consequently, documenting these changes has become a central objective in ecological research, as they offer early indicators of ecosystem disruptions. However, inferring phenological shifts from long-term programs can poses significant methodological challenges (Brown et al., 2011, 2016). On the one hand, direct methods that follow individual parents allow assessing the exact date of the breeding onset (e.g., the laying date) and reproductive investment/output (e.g., the number of laid eggs) for each couple, providing detailed insight into breeding phenology. However, these methods are very labor-intensive, requiring e.g., weekly checks to all nest boxes (for hole-nesting birds) during the entire breeding season. On the other hand, indirect methods that simply count individuals of different classes (e.g., juveniles, mature, …) at the population level, and outside of their nests, are easier to implement, as they only require setting up a few sampling sessions. However they require some post-processing to yield reproductive parameters, and may be subject to biases associated with difficulties in taking into account varying sampling effort (Schmeller et al., 2009).

One simple way to infer breeding onset from population-level monitoring is to track temporal changes in the proportion of juveniles in the population (Fig. 1). By conducting several sampling sessions throughout the breeding season, often with the support of citizen science, researchers obtain a time series of juvenile counts (and proportions) that reflects the underlying distribution of breeding onset. At the start of the breeding season, the proportion of juveniles is zero. It then increases before reaching a plateau after all reproduction is completed (Fig. 1). Clearly, this pattern captures several important aspects of breeding phenology: the position of the curve should change with the mean breeding phenology, its steepness relates to phenological variation (with more variable populations displaying more shallow change), and its asymptote (height of the plateau) should increase with reproductive success. All these

aspects have ecological importance: variance in phenology is a pre-requisite to its heritability and thus evolution under climate change, and reproductive success is more directly relevant to the persistence of a population than breeding time *per se*. However, beyond these qualitative statements, there is surprisingly little quantitative investigation of how well reproductive parameters can be inferred from time series of proportions of juveniles. Moussus et al. (2010) evaluated and proposed different methods to estimate breeding phenology based on counts and age-ratios over the breeding season, and showed that Generalized Additive Models (GAMs) with a fixed shape but shifting position accurately estimated between-year variation in mean fledging phenology. However, they did not take into account phenological variance and reproductive success, which are not only interesting *per se*, but also tightly linked to the estimate of mean phenology, as we show below. More recently, Cuchot et al. (2024, 2025) proposed a parametric method estimating the inflection point, steepness, and asymptote of a sigmoid (logistic) curve for the change in juvenile proportion over the breeding season. This should allow inferring the different aspects of reproductive phenology highlighted above (mean and variance of fledging time, and reproductive success), but this link was left implicit in this approach. There is therefore a need for an approach that explicitly relates reproductive parameters to the proportion of juveniles over time.

An additional limit of methods that rely on observations or captures outside the nest is that they can only estimate the parameters of juvenile emergence, rather than birth. There is generally a time lag before juveniles can be observed in their habitat. For hole-nesting birds, this lag includes the incubation time, plus the time it takes for nestling to be able to fledge and leave the nest. Not only does this lag need to be accounted for to estimate breeding phenology, but it can also lead to significant modifications of the distribution of breeding times. For example in birds, laying date can be subject to strong selection pressures (via the number of eggs and/or fledglings produced), which vary in intensity and timing depending on

environmental conditions (de Villemereuil et al., 2020; Porlier et al., 2012; Visser & Both, 2005). Consequently, the timing of emergence (i.e., rise in frequency) of juveniles outside of their nests, as well as their number, are influenced not only by their birth date and number, but also by the selection they have undergone. This can lead to potential biases when inferring breeding phenology.

In this study, we develop a mathematical model that starts from population-level processes—distribution of laying dates, reproductive success, and selection—to predict the observed proportion of juveniles in samplings over the breeding season. Analytical investigation of this model reveals how (i) the parameters of fledging phenology quantitatively affect the observed proportions of juveniles; and (ii) parameters of breeding phenology can be inferred from fledging phenology, if information about selection is available. We then simulate the process of collecting population-monitoring data throughout the breeding season, and estimating fledging and laying date phenology and reproductive success using this model, for populations with varying degrees of variance and selection pressure. Finally, we compare different sampling designs with varying levels of effort to determine their effectiveness in estimating reproductive parameters across sites.

Methods

Theoretical predictions

General model for proportions of juvenile

The direct observation is number of sampled juveniles J_t and adults A_t at a given site at time *t*. From this we want to infer (1) the distribution $g(z)$ of breeding time z, and (2) the mean reproductive success *R*.

The number J_t of sampled juveniles at time *t* is assumed to be drawn from a binomial distribution, with parameters p_t (the true proportion of juveniles in the population at time *t*) and $N_t = J_t + A_t$ (the total number of sampled individuals at time *t*). Denoting with non-capitalized letters j_t and a_t the true numbers of juveniles and adults in the population at time *t*, we have

$$p_t = \frac{j_t c_{j,t}}{j_t c_{j,t} + a_t c_{a,t}} \quad (1)$$

where $c_{j,t}$ and $c_{a,t}$ are the sampling probabilities of juveniles and adults at time *t*. The critical step in the derivation involves expressing the number of juveniles at time *t* from the cumulative reproduction of parents breeding at different times before, weighted by offspring survival to fledging (and non-dispersal),

$$j_t = \frac{1}{2}\int_{-\infty}^{t-T_f} a_z g(z) W(z) dz. \quad (2)$$

(Note that the integral runs from $-\infty$ for technical convenience but this will not impact the results as long as all individuals reproduce later than time 1 - January 1st - which will be the case in most studies). Importantly, eq. (2) introduces the distribution $g(z)$ of breeding times (e.g. laying dates) in the population, which quantifies the reproductive phenology that we aim to infer. The time to fledge T_f is the time interval between egg laying and fledging of the

juvenile out of the nest, where it may be sampled. The factor half in eq. (2) accounts for the fact that only females produce offspring (assuming even parental sex ratio). The distribution of breeding times $g(z)$ is weighted in eq. (2) by the fitness function $W(z)$, which determines how the time z at which parents breed influences their breeding success, as well as the survival-to-fledging of their hatchlings. Expanding further, we have $W(z) = F(z)S(z)$, where $F(z)$ is the fertility (clutch size) of parents reproducing at time z, and $S(z)$ is the survival probability of individuals hatched from eggs laid at time z. ($S(z)$ could also include the probability of non-dispersal of these offspring, in which case $W(z)$ would not be a fitness function *stricto sensu*). Adults that do not breed also contribute to $W(z)$, simply leading to a reduction of the mean breeding success.

At the end of the breeding season, the total number of fledglings produced is

$$j_\infty = \int_{-\infty}^{\infty} \frac{a_z}{2} g(z) W(z) dz = \frac{\bar{a}}{2} R \quad (3)$$

where $\bar{a} = \int_{-\infty}^{\infty} a_z g(z) dz$ is the mean number of breeding adults over the breeding season (and $\bar{a}/2$ the mean number of breeding females, assuming even sex ratio), and

$$R = \frac{\int_{-\infty}^{\infty} a_z g(z) W(z) dz}{\int_{-\infty}^{\infty} a_z g(z) dx} \quad (4)$$

is the mean reproductive success per breeding pair (or per female), that is, the average number of fledglings they produce over the breeding season. Combining with eq. (1), the asymptotic proportion of juveniles in the population is thus

$$p_\infty = \frac{R}{R + 2r_c} \quad (5)$$

which only depends on the mean reproductive success R and on $r_c = \frac{a_\infty}{\bar{a}} \frac{c_{a,\infty}}{c_{j,\infty}}$, the ratio of the final to the average number of adults over the breeding season, multiplied by the final ratio of sampling probabilities between juveniles and adults.

If we further assume that (i) sampling probabilities are the same in juveniles and adults at any time (but may still vary over time, for instance because of variable sampling effort), such

that $c_{j,t} = c_{a,t} = c_t$; and (ii) the total number of parents in the population does not vary much over a breeding season, such that $a_t \approx a$ for all *t*, then we get

$$p_t = \frac{RG^*(z-T_f)}{2+RG^*(z-T_f)} \quad (6)$$

where the mean reproductive success becomes

$$R = \int_{-\infty}^{\infty} g(z)W(z)dz, \quad (7)$$

and

$$G^*(z) = \int_{-\infty}^{t} \frac{g(z)W(z)}{R} dz \quad (8)$$

is the cumulative distribution of breeding times after selection in the ongoing generation. In other words, the function $G^*(z - T_f)$ is the cumulative probability of successful breeding attempts producing offspring that eventually survive to fledge some time T_f later, allowing them to be potentially sampled as juveniles at the sampling site. The asymptotic proportion of juveniles at the end of the breeding season is

$$p_\infty = \frac{R}{2+R} \quad (9)$$

Reciprocally, assuming that the curve relating the proportion p_t of juveniles to time has been fitted empirically, this curve can be used to estimate the mean reproductive success as

$$\hat{R} = \frac{2p_\infty}{1-p_\infty}, \quad (10)$$

and the cumulative distribution (after selection) of breeding times up to time *z* as (combining eqs. (6) and (9))

$$\widehat{G^*}(z - T_f) = \frac{p_z\,(1-p_\infty)}{(1-p_z)\,p_\infty}. \quad (11)$$

Equation (11) thus shows that the cumulative distribution of breeding times, accounting for potential effects of selection in the ongoing generation, is simply the odds ratio of frequency of juveniles at time *z* over their final frequency. The distribution of fledging times (rather than

breeding times) is simply obtained by removing the term $-T_f$ in eq. (11). We will now proceed to derive more explicit results based on this foundation.

Analytical predictions under specific assumptions

Logistic curve for p_t

For practical purposes, it is convenient to fit the proportion of juveniles in catches as a logistic function of time,

$$p_t = \frac{p_\infty}{1+\exp\left(-\frac{t-t_m}{b}\right)} \quad (12)$$

where p_∞ is the asymptotic (maximum) proportion of juveniles (as in eq. 9), t_m is the time at which $p_t = p_\infty/2$ (midpoint and inflection point of the curve), and b controls the slope at t_m (with steeper slopes under smaller b). What makes equation (12) especially useful is that it can be combined with binomial error to yield a modified logistic regression, allowing estimation of p_t from raw counts J_t and A_t for juveniles and adults, respectively (Cuchot et al., 2024). Combining with eq. (11), taking the derivative with respect to time, and scaling to an integral of 1, leads to the probability density function of breeding times, from which all moments of the distribution can be derived. In particular, the estimated mean breeding time after selection is

$$\widehat{\mu^*} = t_m - T_f - b\log(1-p_\infty) \quad (13)$$

(the mean *fledging* time is simply $\widehat{\mu^*} + T_f = t_m - b\log(1-p_\infty)$), and the variance in fledging time (or breeding time weighted by selection) is

$$\widehat{\sigma_z^{2*}} = \frac{\pi^2 b^2}{3}. \quad (14)$$

Interestingly, eq. (13) shows that the mean breeding time in the population is not directly estimated by the midpoint t_m of the logistic curve for p_t (corrected by fledging time T_f). Instead, the true mean breeding time occurs later (since $-\log(1-p_\infty) > 0$ for all $0 < p_\infty < 1$), all the more so as the slope of the logistic function is shallower (larger b), and the final

196 frequency of juveniles p_∞ is larger. In other words, the midpoint of the frequency of juveniles
197 in catches precedes the true mean fledging time, with a larger advance when the mean
198 reproductive success (from eq. 9) and/or the variance of breeding time (eq. 14) are large (Fig.
199 1). On the other hand, the variance of breeding times is proportional to the squared slope
200 parameter b^2, and does not depend on any other parameters of the logistic curve.

201 Gaussian fitness peak

202 As mentioned above, whenever breeding time influences fitness by affecting clutch size and/or
203 survival to fledging, samples outside of the nests can only infer the post-selection distribution
204 of breeding time (density $g^*(z)$ and cumulative distribution $G^*(z)$), not the baseline
205 distribution $g(z)$ of breeding times in parents before selection. To understand how these two
206 distributions are related, we assume that the pre-selection distribution $g(z)$ is normal, with
207 mean μ and variance σ_z^2, and make use of classic results from evolutionary quantitative genetics
208 (Lande, 1976; Walsh et al., 2018). In addition, more analytical progress can be achieved when
209 also assuming a specific shape for the fitness function $W(z) = F(z)S(z)$. A typical assumption
210 for phenology, supported by empirical analyses (Chevin et al., 2015; de Villemereuil et al.,
211 2020; Gamelon et al., 2018; Reed et al., 2013), is that fitness is maximized at some intermediate
212 optimum breeding time z, such that reproducing too early or too late is detrimental. This is
213 typically modelled by making fitness a Gaussian function of breeding time,

214 $$W(z) = W_{\max} \exp\left(-\frac{(z-\theta)^2}{2\omega^2}\right), \quad (15)$$

215 where θ is the optimal breeding time where reproductive success is highest (fitness peak), $W_{\max}$
216 is the maximum fitness of parents that breed at that optimal timing (peak height), and ω is the
217 width of the fitness peak, with narrower peaks leading to stronger stabilizing selection. If both
218 fertility and viability exert stabilizing selection in a form similar to eq. (15), then the effective

width, height, and optimum of the peak for the resulting total fitness function in eq. (15) can be derived from their basic parameters (Cotto & Chevin, 2020).

Under these assumptions, the mean breeding time after selection in the ongoing generation is (modified from Lande, 1976)

$$\mu^* = \mu - \frac{\delta}{\tilde{\omega}^2+1}, \quad (16)$$

where $\delta = \mu - \theta$ is the mean phenotypic mismatch with optimum (positive for delayed reproduction), and $\tilde{\omega} = \omega/\sigma_z$ is the standardized peak width (scaled to the standard deviation of the trait). Equation (16) shows that selection in the ongoing generation changes the mean phenotype in a direction opposite to the phenotypic mismatch with optimum δ (hence the minus sign), thus bringing the mean phenotype back towards the optimum, to an extent that increases with the magnitude of the mismatch δ and the standardized strength of stabilizing selection (stronger with smaller $\tilde{\omega}$) (Lande, 1976). The phenotypic variance after selection is

$$\sigma_z^{2*} = \frac{\tilde{\omega}^2}{\tilde{\omega}^2+1}\sigma_z^2 \quad (17)$$

which importantly does not depend on the mismatch δ. The ratio in eq. (17) is always lower than 1, showing that selection for an optimum phenotype always reduces phenotypic variance, all the more so as peak width is small relative to the phenotypic variance (small $\tilde{\omega}$), leading to strong stabilizing selection. Stabilizing selection also influences the mean reproductive success R, which is equal to the mean fitness (eq. 7), such that

$$\hat{R} = W_{\max}\sqrt{\frac{\tilde{\omega}^2}{\tilde{\omega}^2+1}}\exp\left(-\frac{\tilde{\delta}^2}{2(\tilde{\omega}^2+1)}\right), \quad (18)$$

where $\tilde{\delta} = \delta/\sigma_z$ is the standardized phenotypic mismatch. Equation (18) shows that for a given reproductive potential (quantified by $W_{\max}$), the mean reproductive success is reduced when the mean phenotype deviates from the optimum ($\tilde{\delta} \neq 0$), as expected. But even when the mean phenotype *is* at the optimum, the reproductive output is still reduced by phenotypic variance

causing some individuals to deviate from the optimum (decreasing $\sqrt{\frac{\tilde{\omega}^2}{\tilde{\omega}^2+1}} \leq 1$, variance load, Lande & Shannon, 1996).

If the standardized peak width $\tilde{\omega}$ and mismatch with optimum (both raw δ and standardized $\tilde{\delta}$) can be known independently, then it is possible to work backwards from post-selection estimates based on the proportion of juveniles over time (eqs (10), (13) and (14)), combined with the effect of selection in eqs (16-18), to infer pre-selection breeding parameters as

$$\hat{\mu} = \hat{\mu}^* + \frac{\delta}{\tilde{\omega}^2 + 1}$$

$$\widehat{\sigma_z^2} = \frac{\tilde{\omega}^2+1}{\tilde{\omega}^2} \widehat{\sigma_z^{2*}} \quad (19)$$

$$\widehat{W_{\max}} = \hat{R} \sqrt{\frac{\tilde{\omega}^2 + 1}{\tilde{\omega}^2}} \exp\left(\frac{\tilde{\delta}^2}{2(\tilde{\omega}^2 + 1)}\right)$$

where the hat notation denotes a statistical estimate.

Simulations

Validating analytical predictions

To validate the analytical results above and assess their ability to estimate the true breeding phenology, we simulated datasets of breeding time and sampling of juveniles and parents, varying $W_{\max}$, σ and ω. We drew the laying dates of 1000 breeding pairs from a normal distribution with mean μ, fixed to 90 in the simulations, and standard deviation σ, ranging from 3 to 10. The number of sampling sessions was initially set to 150 to be non-limiting (but see below). We set the time to fledge T_f to 40, as this is generally considered to be a good approximation for blue and great tits (20 day of incubation followed by about 20 days to fledge). We modeled the relationship between breeding time and reproductive success for each

breeding pair using a Gaussian fitness function (following eq. 15), centered on an optimum θ set to 20 days before mean laying date. This 20-day delay was introduced to account for the fact that selection for earlier breeding is generally found in birds, and is often ascribed to a mismatch with the optimal laying date (Chevin et al., 2015; de Villemereuil et al., 2020; Gamelon et al., 2018; Reed et al., 2013). The Gaussian fitness peak is also characterized by its standardized width $\tilde{\omega}$ determining the intensity of stabilizing selection, with smaller $\tilde{\omega}$ leading to a narrower peak and thus stronger stabilizing selection. The peak height $W_{\max}$ sets the reproductive success of well-adapted parents with the optimal laying date. The number of eggs laid by each pair was drawn from a Poisson distribution, with expectation provided by the fitness function evaluated at their laying date. Once the laying dates and the number of eggs per pair were established, we simulated sampling of juveniles and adults from the population at specific times corresponding to sampling sessions, evenly spaced between the start and end of the sampling protocol.

Each parameter combination was simulated 10 times, and the results were averaged over these replicates. For each combination of parameters and simulation repeat, we estimated a sigmoid curve for the proportion of juveniles over time from the population-monitoring data, implementing the non-linear sigmoid model from equation (12) using R2jags (Su & Yajima, 2021) package on R (R Core Team, 2024). We ran the models with 4000 iterations, 1000 of burning, 3 chains and with weakly informative priors (Appendix 3). We then transformed estimates of parameters from this function to obtain the required fledging parameters from equations (10), (13) and (14). We also predicted the mean and variance of breeding time, (before selection) and the maximum fitness W_{max}, from the (assumed to known) selection parameters (equation 19).

Comparing sampling schemes

In the simulations described in the previous section, we deliberately set the numbers of sampling sessions and breeding pairs to very high values, enabling us to establish clear links between breeding and selection parameters on one hand, and proportion of sampled juveniles on the other hand, regardless of considerations on inference strength. However, this was not meant to realistically model real sampling schemes, which generally include only a handful of sampling sessions. Fewer sampling sessions and number of breeding pairs, which jointly characterize sampling effort, are expected to reduce the accuracy of the estimates of breeding parameters (Arizaga et al., 2023). Nevertheless, Cuchot et al. (2024) suggested that the lack of annual data per site could be compensated by pooling data from multiple sites, including year and site as random effects in the analysis.

To assess this claim, we next simulated data with more realistic parameter values, corresponding to two typical designs from European ringing program (Robinson, 2023), differing by their number of capture (i.e. sampling) sessions per year (from three to ten). For each set of simulated data, we proceeded as if we had pooled 150 populations that varied in their mean laying dates, but otherwise had similar distribution of phenology (normal distributions with the same within-population variance). Grand mean laying date was set to 90 (as previously) and ω (the unstandardized peak width) to 100 (vanishingly weak selection). The between-site standard deviation in laying dates was set to 10, and we compared different magnitudes of within-population standard deviations (1, 5 and 10). This approach allowed us to assess how the sampling design (number of capture sessions) and sampling effort (number of capturable pairs on the sampling site, set to 15 for low sampling effort and 200 for high sampling effort) affects the accuracy and precision of estimates for breeding phenology. It also allowed to assess how well the between-site variance in laying date could be estimated. During the sampling sessions, we set the total number of sampled individuals to be 80% of the flying

311 individuals (include the parents and fledglings). For each set of parameters, we fit the same
312 non-linear model (eq. 12) to these simulated datasets, but now also including a random effect
313 for site on the sigmoid inflection point t_m(equation 20, 21).

314 $$p_{t,site} = \frac{p_\infty}{1+\exp\left(-\frac{t-(t_m+\rho_{site})}{b}\right)} \quad (20)$$

315

316 $$\rho_{site} \sim N(\mu_{\rho\ site}, {\sigma^2}_{site}) \quad (21)$$

317 with ${\sigma^2}_{site}$ being the between-site variance. We ran the models with 50000 iterations, 20000
318 of burning. The model is detailed in appendix 3. From this we calculated $\widehat{\mu^*}$ from equations
319 (13), ${\sigma^2}_{site}$ from model output, and compared them to the predictions from simulated
320 parameters, across varying numbers of capture sessions and breeding pairs. Each combination
321 of parameter and model was simulated 20 times.

322

Results

Analytical predictions for breeding phenology

We derived an analytical model to predict observed proportions of sampled juveniles over the breeding seasons from the underlying distribution of laying dates and mean reproductive success in a population. A striking result from this model is that, even without selection, the mean fledging time is not directly predicted by the inflection point (or midpoint) of the curve relating the proportion of juveniles to time. Instead, the mean fledging time lags some time after the midpoint. From eqs (9), (13) and (14), this lag is

$$L = \sqrt{3}\frac{\sigma_z}{\pi}\log\left(\frac{2+R}{2}\right) \quad (22)$$

which increases proportionally to the phenotypic standard deviation in breeding times σ_z, and also increases with the mean reproductive success R. This is illustrated in Figure 1, which shows that the true mean fledging time (dashed gray line, occurring sometime after the mean breeding time shown in continuous gray line) occurs later than the midpoint in the proportion of juveniles, all the more so as phenotypic variance (Fig. 1A) or the reproductive success (Fig. 1B) are large. The reason for this lag L is that the proportion of juveniles in sampled individuals is not the cumulative distribution of fledging times *per se*, as it also depends on the number of adults. The midpoint therefore needs to be transformed using eq. (11) to reach information about fledging times (or breeding time, after correcting by time to fledge).

An important consequence of eq. (22) is that the mean reproductive success and the variance in breeding times, even when they are not the main focus *per se*, are both key to predicting the mean phenology. Fortunately, these parameters can also be estimated from proportion of juveniles among samples using simple formulas (eqs. 10 and 14, respectively). Hence, the analytical theory predicts that the sigmoid curve relating the proportion of juveniles

in samples to time contains important quantitative information about reproductive phenology and success. As a proof of principle, and since some of these predictions were based on mathematical approximations, we have verified them using simulations.

Estimating fledging parameters in simulations.

Our analyses of simulation results demonstrate that the method yields accurate estimates of fledging phenology and reproductive success (Fig. 2). The mean fledging time is well estimated by eq. (13) (solid vs dashed black lines in Fig. 2A), while directly using the midpoint of the logistic curve leads to anticipating the true fledging time by as much as 10 days in our example (grey lines in Fig. 2A). This delay increases with higher variance in breeding time and reproductive success, as predicted by eq. (22). Our approach also correctly estimates the variance in fledging time, although underestimating it slightly (Fig. 2B). This underestimation is partly explained by the fact that the realized variance in the population is expected to be lower than the theoretical variance because of the finite number of breeding pairs (reduced in proportion $1 - 1/n$, with n the number of pairs). This underestimation of the phenological variance leads to a small underestimation of the mean fledging time (continuous vs dashed lines in Fig. 2A), owing to the dependence of the estimated mean fledging time on the variance (eqs. 13 and 22). The asymptotic proportion of juveniles consistently provides robust estimates of the reproductive success (Fig. 2C).

From fledging to breeding parameters.

Having validated that the fledging phenology and reproductive success can be estimated accurately through sampled juvenile, we next proceed to inferring phenological parameters prior to fledging, relating to the laying date of breeding pairs. Beyond just accounting for the delay between the laying date and fledging date quantified by the time to fledge T_f, we also

allow for selection, by letting the reproductive success of a breeding pair (number of offspring that reach the fledgling stage) depend on the match between their laying date and an optimum laying date that is earlier than the mean phenotype in the population (following eq. 15, consistent with Chevin et al., 2015; de Villemereuil et al., 2020). We consider estimates that do, or do not, correct for the influence of selection, assuming in the former case that selection parameters can be estimated through other means.

For a given mean laying date (dashed grey line in Fig. 3A, fixed at 90 days), selection for earlier breeding causes the mean fledging time to become earlier, all the more so as the phenotypic variance is large (dashed black line in Fig. 3A) and selection is strong (smaller ω, lower panel in Fig. 3A). If selection is not accounted for and the mean laying date is simply estimated by subtracting T_f from the mean fledging time, this leads to a systematic downward bias in the estimation of laying dates (light grey line in Fig. 3A). The magnitude of this bias increases with greater variance in laying date and stronger selection, mirroring the trend in the mean fledging time. Stabilizing selection induced by the fitness peak with an optimum also significantly reduces the variance in fledging dates (black dashed lines in Fig. 3B) relative to the variance in laying dates (grey dashed lines in Fig. 3B), all the more so as selection is stronger (lower panel in Fig. 3B). Finally, selection causes the realized mean reproductive success in the population (black dashed lines in Fig. 3C) to be lower than the maximum reproductive success of a pair with optimal laying date (grey dashed lines in Fig. 3C), all the more so as selection is strong. These influences of selection on phenological and reproductive parameters qualitatively match the predictions from eqs. (16-18).

Assuming that the parameters of the fitness function in eq. (15) are known from other approaches, we are able to correct for the influence of selection on all estimated breeding parameters, working backwards from the fledging parameters estimated through the proportion of juveniles in samples. The mean laying date (Fig. 3A) and maximum reproductive success

394 are very well estimated in all conditions (continuous vs dashed lines in Fig. 3C; see also
395 Appendix 1, Figs. S1-3). The standard deviation of breeding phenology is still underestimated
396 (dashed vs continuous line of same color in Fig. 3B), as without selection (Fig. 2B), but
397 interestingly this bias seems to be less pronounced as selection becomes stronger (lower vs
398 upper row in Fig. 3B; see also Appendix 1, Fig. S2). Furthermore, the effect of selection on
399 phenological SD is well corrected by our approach (black vs grey lines in Fig. 3B).

400 ## Power in a multi-site approach

401 So far, we have used ideal conditions as proof of principle for the validity of our approach. We
402 now investigate the method's performance under more realistic conditions. Specifically, we
403 focus on scenarios involving pooled data from multiple sites, each with moderate sampling
404 effort, and variation in mean laying date between sites. We incorporate random effects in the
405 midpoint of the sigmoid curve to improve estimation of grand mean phenology (across sites),
406 together with providing information about site-to-site variation in mean phenology.

407 For a given mean laying date (dashed black line in Fig. 4A, fixed at 90 days), large
408 within-site variance in laying date causes mean fledging time to be estimated earlier (from 1 to
409 3 days difference, depending on the number of sampling session) when sampling effort is low
410 (Fig. 4A). Within-site variation in laying date is systematically underestimated, but this bias is
411 reduced when switching to high sampling effort and increasing the number of capture sessions
412 from 4 to 10. The reproductive success is systematically very well estimated, with in the worst
413 case a difference of 0.5 offspring with the true value, with slight under-estimation in the
414 scenarios with high within-site variation in laying date (Fig. 4C). Between-site variation in
415 fledging date is slightly underestimated, except in scenarios with the fewest sampling sessions
416 (n = 3), where it is slightly overestimated—particularly in scenarios with low variation in
417 fledging time (Fig. S4). As described in eq. (13), mean and variance in fledging time are linked,

418 and this relationship may explain why the mean fledging time underestimated in scenarios
419 where the standard deviation is also underestimated (Fig. 4A-B). These simulations highlight
420 the trade-offs between maximizing the number of sessions, sites, or individuals sampled in
421 designing effective sampling schemes.

Discussion

Assessing reproductive parameters from population samples over time

In this study, we aimed at understanding how to infer breeding phenology - laying and fledging time - and reproductive success from the emergence of juveniles among sampled individuals, typically passerines in mist-net capture schemes. Our approach confirmed that a three-parameter function describing how the proportion of juveniles in samples changes along the breeding seasons allows a good assessment of the mean population breeding phenology. However, we show that the midpoint of the sigmoid curve does not directly predict this mean phenology. Instead, it needs to be corrected using the asymptote and maximum slope of the sigmoid curve (eq. 13). Failing to do so may lead to inferring spurious correlations between the mean phenology and the phenological variance, or the mean reproductive success. Nevertheless, if only the midpoint varies, it can still be used to study variation in mean phenology across sites and time (Fig. 4 and Fig. S2). Moreover, we showed that, beyond just the average breeding phenology, this method also allows a good assessment of phenological variance and the reproductive success, respectively from the maximum slope and asymptote of the curve. Finally, we demonstrated that this method can leverage hierarchical modeling to estimate breeding parameters across sampling sites. We also identified important aspects of the sampling design, suggesting that while too few sampling sessions should be avoided, increasing their number does not improve the accuracy or correctness of the estimates, as long as many sites are available.

Accounting for selection

Although the method does not allow for direct estimation of the actual strength of selection, it provides valuable insights into how selection on laying date, via clutch size and nestling

survival to fledging, can influence our estimates of breeding phenology. Specifically, strong selection for earlier breeding introduces a systematic bias in the laying dates inferred from sampled juveniles, which increases with variance in laying dates (light grey line in Fig. 3). When parameters of selection are not known, equations (10), (13) and (14) can still correctly estimate the post-selection breeding parameters (fledging phenology), and the (co)variation in these estimates across years and sites may possibly provide information about variation in selection. For instance, equations (16) and (18) predict that years of strong maladaptation lead to joint changes in the mean breeding time after selection $\bar{z}^*$ and reproductive output R, through their shared dependence on the mismatch with optimum δ.

Furthermore, if the basic parameters of the fitness function can be known from other sources, then we show how to efficiently correct for the influence of selection, when working back from fledging to laying date parameters (dark gray lines in Fig. 3). These parameters of the fitness function could be estimated from a small subset of sites where both population (e.g., mist-net) and nest-box monitoring are available. Where this is not feasible, estimates from the literature (Chevin et al., 2015; de Villemereuil et al., 2020) could be used to put bounds on the possible influence of selection on the inferred breeding phenology, conditional on the estimated fledging phenology.

Sampling effort

European monitoring program of bird capture widely vary in both the number of sampling sessions, ranging from to 3 to 10, and in the sampling effort per session, i.e., the number of mist net settled (Robinson, 2023). Here, we showed that sampling only three times during the breeding season led to overestimating the average fledging phenology. However, it should be noted that in the worst-case scenario, the estimated fledging time was delayed by only three days. Increasing the sampling effort (having access to more breeders) and the number of

sampling sessions affected the estimation of the within-site SD in fledging time (Fig. 4), but the latter was systematically underestimated.

Based on these results, for a typical constant ringing monitoring program of bird populations, we recommend conducting between four and five sampling sessions during the breeding season with moderate effort. Fewer than four sessions may not capture enough variability in phenological parameters, while increasing the number of sessions beyond five does not improve the precision or accuracy of the estimates. This approach strikes an optimal balance between resource investment and the reliability of the monitoring results. Crucially, this would allow increasing the number of sites where these sampling sessions are performed, which is critical to the accuracy of the inference.

Sampling probabilities

Our model assumes an equal sampling probability between age status (adult or juvenile). In reality, juveniles might be easier to capture due to their lower flight abilities (Senar et al., 1999). While this statement seems broadly valid, evidence suggests that it depends on the sampled species. For instance, no difference was found in house sparrows (Senar et al., 1999), while higher capture probability of juveniles was found in starlings (Simons et al., 2015).Variation in the activity of the parents over time could also modify their chance of being captured. Presumably they should be more active before fledging, because they need to provide all the food to juveniles. After fledging, the juveniles take care of themselves, so their parent's workload must decline quite dramatically. Such biases could affect the shape of the logistic curve on which our analysis relies, for instance, not leading to a plateau (i.e., asymptote) in the case where adults capture probability decrease with time. However, if information about these parameters, such as changes in capture probability, is available, it could be directly incorporated into the analysis using eq. 1 to account for such effects.

Conclusion

In the face of climate change, shifts in the timing of biological events in wild organisms have become increasingly significant. Recognizing the need for reliable methods to track these changes, our study demonstrates that a logistic sigmoid curve can accurately estimate breeding phenology and reproductive success from mist-net captures in passerines (and similar population-monitoring designs), provided the parameters of this curve are suitably transformed. In addition, the influence of selection can be accounted for if the fitness function on laying date is available from other approaches. While limitations such as sampling biases remain, this method suggest that an optimized sampling design of four to five sessions yields precise and reliable estimates. Ultimately, this work not only refines methods for avian reproductive monitoring, but also provides a valuable framework for tracking climate-driven phenological changes in wildlife populations, including by relying on citizen size.

References

Arizaga, J., Crespo, A., & Iraeta, A. (2023). Lowering the cost of citizen science: Can we reduce the number of sampling visits in a constant ringing effort-based monitoring program? *Journal of Ornithology*, *164*(1), 245–251. https://doi.org/10.1007/s10336-022-02019-7

Chevin, L.-M., Visser, M. E., & Tufto, J. (2015). Estimating the variation, autocorrelation, and environmental sensitivity of phenotypic selection: ESTIMATING FLUCTUATING SELECTION. *Evolution*, *69*(9), 2319–2332. https://doi.org/10.1111/evo.12741

Cotto, O., & Chevin, L.-M. (2020). Fluctuations in lifetime selection in an autocorrelated environment. *Theoretical Population Biology*, *134*, 119–128. https://doi.org/10.1016/j.tpb.2020.03.002

Cuchot, P., Bonnet, T., Dehorter, O., Henry, P.-Y., & Teplitsky, C. (2024). How interacting anthropogenic pressures alter the plasticity of breeding time in two common songbirds. *Journal of Animal Ecology*, *93*(7), 918–931. https://doi.org/10.1111/1365-2656.14113

de Villemereuil, P., Charmantier, A., Arlt, D., Bize, P., Brekke, P., Brouwer, L., Cockburn, A., Côté, S. D., Dobson, F. S., Evans, S. R., Festa-Bianchet, M., Gamelon, M., Hamel, S., Hegelbach, J., Jerstad, K., Kempenaers, B., Kruuk, L. E. B., Kumpula, J., Kvalnes, T., … Chevin, L.-M. (2020). Fluctuating optimum and temporally variable selection on breeding date in birds and mammals. *Proceedings of the National Academy of Sciences*, *117*(50), 31969–31978. https://doi.org/10.1073/pnas.2009003117

Gamelon, M., Tufto, J., Nilsson, A. L. K., Jerstad, K., Røstad, O. W., Stenseth, N. C., & Sæther, B.-E. (2018). Environmental drivers of varying selective optima in a small passerine: A multivariate, multiepisodic approach. *Evolution*, *72*(11), 2325–2342. https://doi.org/10.1111/evo.13610

Lande, R. (1976). Natural Selection and Random Genetic Drift in Phenotypic Evolution. *Evolution*, *30*(2), 314–334. https://doi.org/10.1111/j.1558-5646.1976.tb00911.x

Lande, R., & Shannon, S. (1996). THE ROLE OF GENETIC VARIATION IN ADAPTATION AND POPULATION PERSISTENCE IN A CHANGING ENVIRONMENT. *Evolution*, *50*(1), 434–437. https://doi.org/10.1111/j.1558-5646.1996.tb04504.x

Porlier, M., Charmantier, A., Bourgault, P., Perret, P., Blondel, J., & Garant, D. (2012). Variation in phenotypic plasticity and selection patterns in blue tit breeding time: Between- and within-population comparisons. *Journal of Animal Ecology*, *81*(5), 1041–1051. https://doi.org/10.1111/j.1365-2656.2012.01996.x

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Reed, T. E., Jenouvrier, S., & Visser, M. E. (2013). Phenological mismatch strongly affects individual fitness but not population demography in a woodland passerine. *Journal of Animal Ecology*, *82*(1), 131–144. https://doi.org/10.1111/j.1365-2656.2012.02020.x

Robinson, R. A. (2023). Understanding population change: The value of the EuroCES constant-effort ringing programme. *Ringing & Migration*, *38*(1–2), 29–37. https://doi.org/10.1080/03078698.2024.2311771

Schmeller, D. S., Henry, P.-Y., Julliard, R., Gruber, B., Clobert, J., Dziock, F., Lengyel, S., Nowicki, P., Déri, E., Budrys, E., Kull, T., Tali, K., Bauch, B., Set^{TEL}e, J., Van Swaay, C., Kobler, A., Babij, V., Papastergiadou, E., & Henle, K. (2009). Advantages of Volunteer-Based Biodiversity Monitoring in Europe. *Conservation Biology*, *23*(2), 307–316. https://doi.org/10.1111/j.1523-1739.2008.01125.x

Senar, J. C., Conroy, M. J., Carrascal, L. M., Domènech, J., Mozetich, I., & Uribe, F. (1999). Identifying sources of heterogeneity in capture probabilities: An example using the Great Tit *Parus major*. *Bird Study*, *46*(sup1), S248–S252. https://doi.org/10.1080/00063659909477251

Simons, M. J. P., Winney, I., Nakagawa, S., Burke, T., & Schroeder, J. (2015). Limited catching bias in a wild population of birds with near-complete census information. *Ecology and Evolution*, *5*(16), 3500–3506. https://doi.org/10.1002/ece3.1623

Su, Y., & Yajima, M. (2021). R2jags: Using R to Run "JAGS." *R Package Version 0.7-1, <https://CRAN.R-Project.Org/package=R2jags>*.

Visser, M. E., & Both, C. (2005). Shifts in phenology due to global climate change: The need for a yardstick. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1581), 2561–2569. https://doi.org/10.1098/rspb.2005.3356

Walsh, B., Lynch, M., & Lynch, M. (2018). *Evolution and selection of quantitative traits*. Oxford University Press.

Author contributions

Paul Cuchot and Luis-Miguel Chevin designed the methodology. L-M.C. conceived the theoretical predictions and P.C. carried out the simulation study. All authors contributed equally to the drafts and gave final approval for publication.

Acknowledgments

We sincerely thank Pierre-Yves Henry and Céline Teplitsky for their invaluable discussions and contributions in initiating this project on fledging phenology. Their insights and support have been instrumental in shaping this work. We also acknowledge the BNP Foundation for their financial support.

Data availability statement

Codes for simulation and statistical model are available in Appendix.

Conflict of interest

Authors declare no conflict of interest.

Figures

Figure 1:

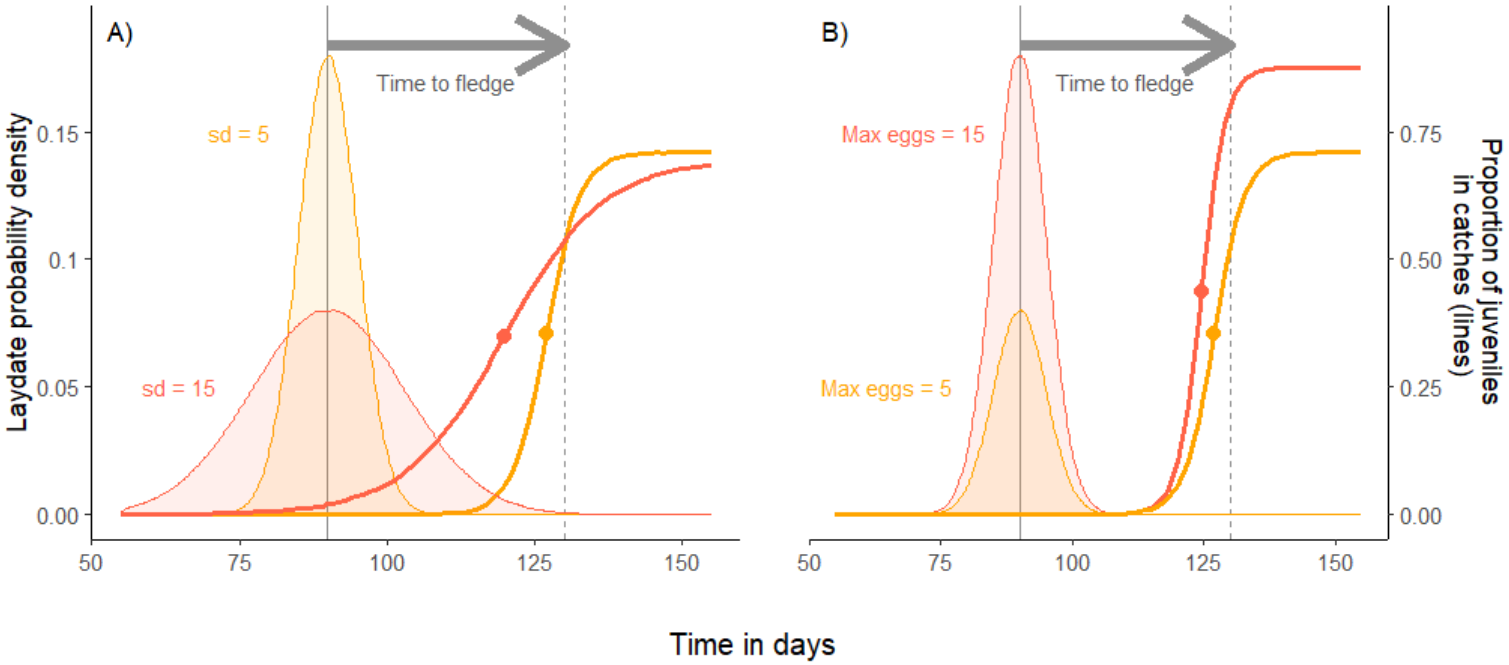
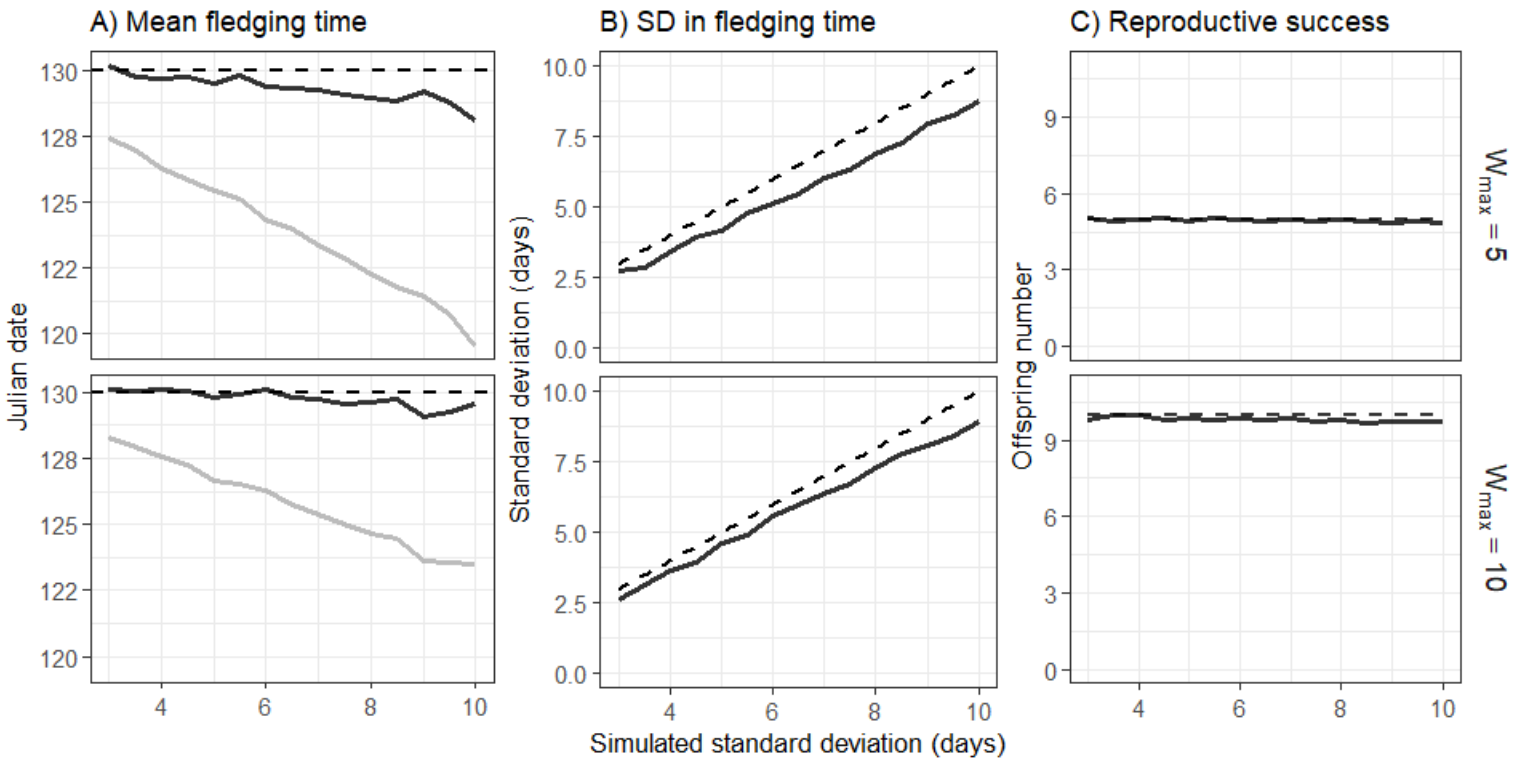


Figure 1: Relationship between breeding phenology and proportion of juveniles in catches. For each panel, the Gaussian curves represent the distribution of laying dates, while the sigmoid curves with the same colors represent the corresponding increases in the proportion of juveniles along the breeding season. For a given mean breeding time (vertical continuous line) and fledging time (vertical dashed line), the inflection point of the sigmoid curve (dot) changes with (A) within-site variation in laying date, and (B) maximal reproductive success (e.g., number of eggs laid). The inflection points are earlier than the true mean fledging time, all the more as variance of fledging time (A) and reproductive success (B) are large (red in both panels). The time to fledge, the time from laying to leaving the nest, is set to 40 days.

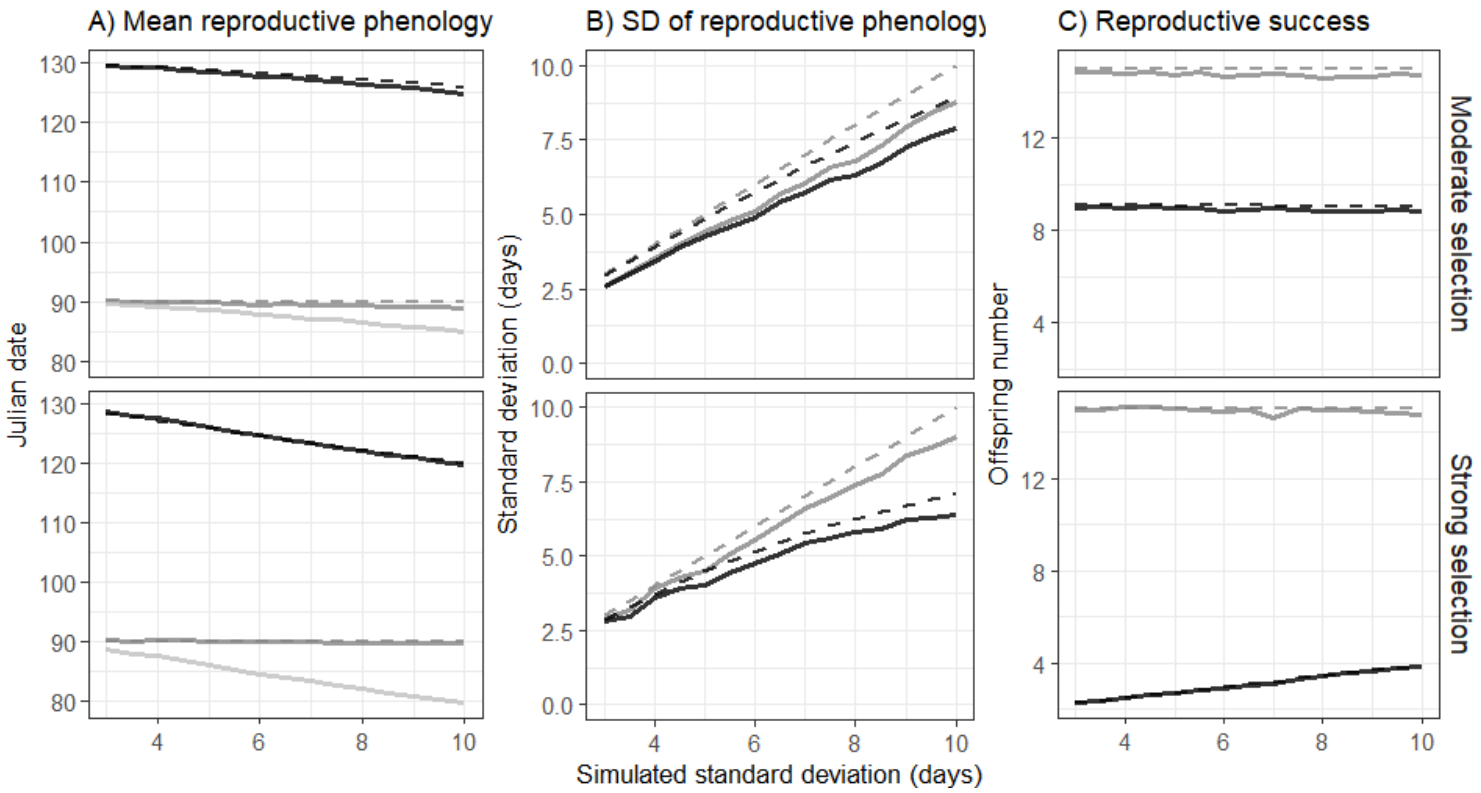
592 Figure 2:



593

594 **Figure 2: Inferring fledging parameters from changes in proportions of juveniles in**
595 **population samples.** The mean fledging time (A), standard deviation in fledging time (B) and
596 reproductive success (C), as estimated from proportions of juveniles in population samples, are
597 shown across a range of true simulated SD of laying date, for two values of maximal
598 reproductive success. The dashed lines represent the known (simulated) values for each
599 parameter, and the solid lines represent estimates from our approach, using eq. (13) in A
600 (representing $\widehat{\mu^*} + T_f$ for the mean fledging time), eq. (14) in B, and eq. (9) in C. The inflection
601 point of the sigmoid (t_m in eq. 12) is also shown as grey line in A.

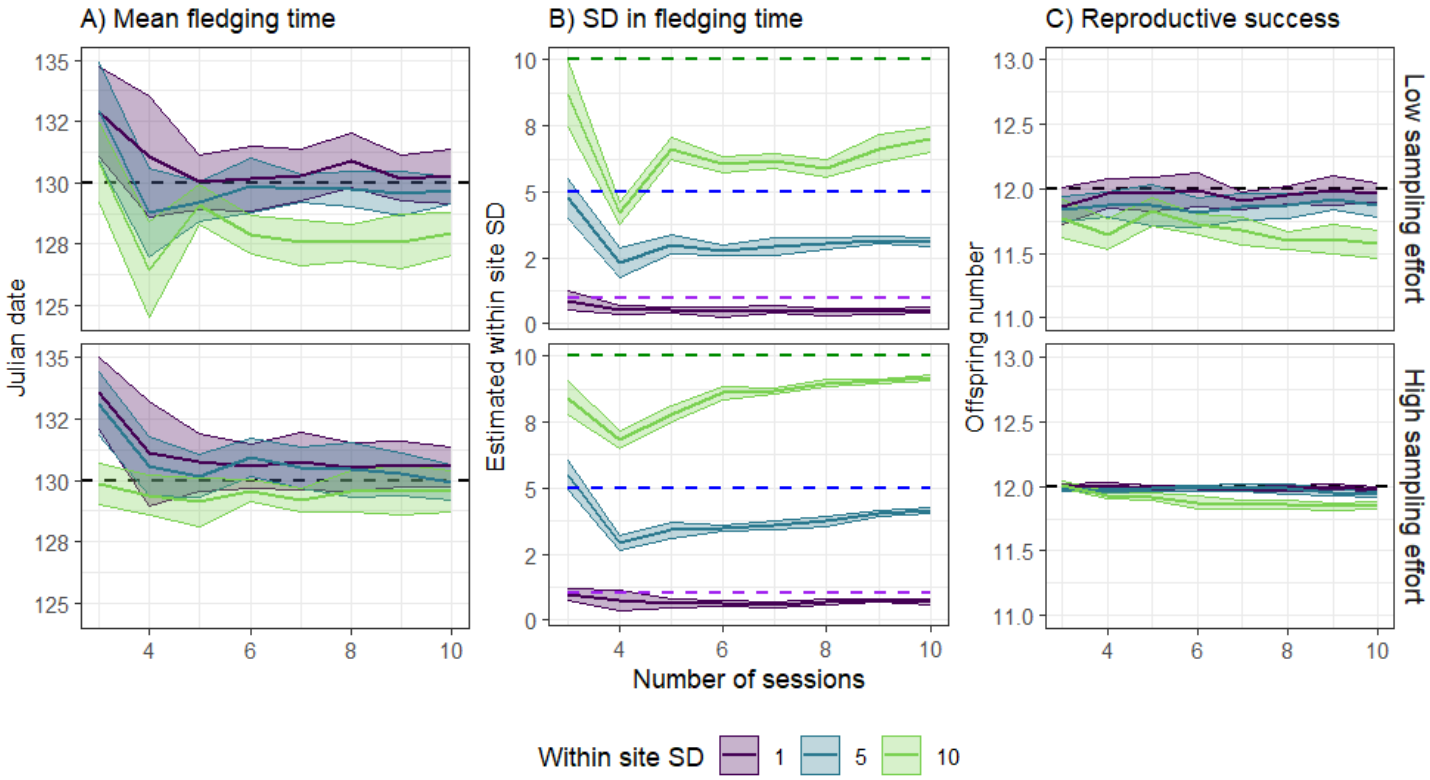
602 Figure 3:



603

604 **Figure 3: Inferring breeding parameters from fledging data**. The mean (A) and standard
605 deviation (B) of reproductive phenology, and the reproductive success (C), as estimated from
606 the proportions of juveniles in population samples (black lines), or inferred for laying (i.e.,
607 breeding) parameters (dark grey lines), are shown across a range of the true simulated SD in
608 laying dates, for two strengths of selection ($\omega = 20$ for "Moderate selection", upper panel;
609 $\omega = 10$ for "Strong selection", lower panel). Dashed lines represent the known (simulated)
610 values for each parameter, while solid lines represent the estimated ones. The light grey solid
611 line in A represents the mean laying date inferred when ignoring the influence of selection, by
612 simply subtracting the time to fledge T_f from the mean fledging date. The dark grey lines in C
613 represent the maximal reproductive success associated to the optimal breeding time.

614 Figure 4:



615

616 **Figure 4: Influence of sampling scheme on cross-site inference of fledging parameters**.
617 The mean fledging time (A), standard deviation in fledging time (B), and reproductive success
618 (C), as estimated from proportions of juveniles in population samples at n = 150 sites, are
619 shown across a range of number of sampling sessions, for two values sampling effort (15
620 capturable breeding pairs for "low sampling effort", upper panel; 200 capturable breeding pairs
621 for "high sampling effort", lower panel) and three values of within-site SD. Selection is set to
622 be weak ($\omega = 100$). Dashed lines represent the known (simulated) values for each parameter,
623 while solid lines represent their mean estimate. Shaded areas correspond to standard deviation
624 of estimated values among the 10 iterations.

625

Supplementary materials

Appendix 1: Validating analytical predictions

Figure S1: Phenology

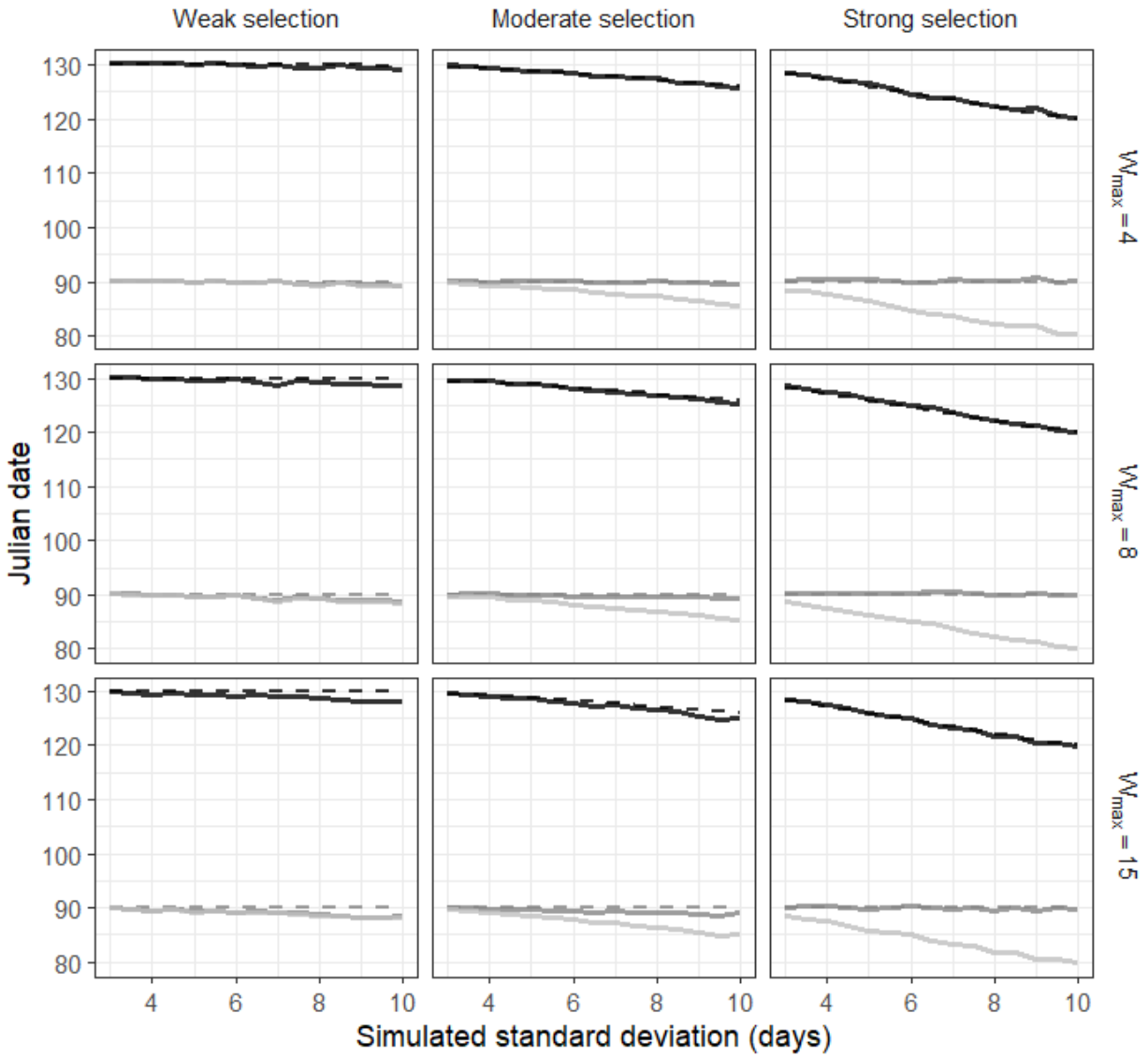
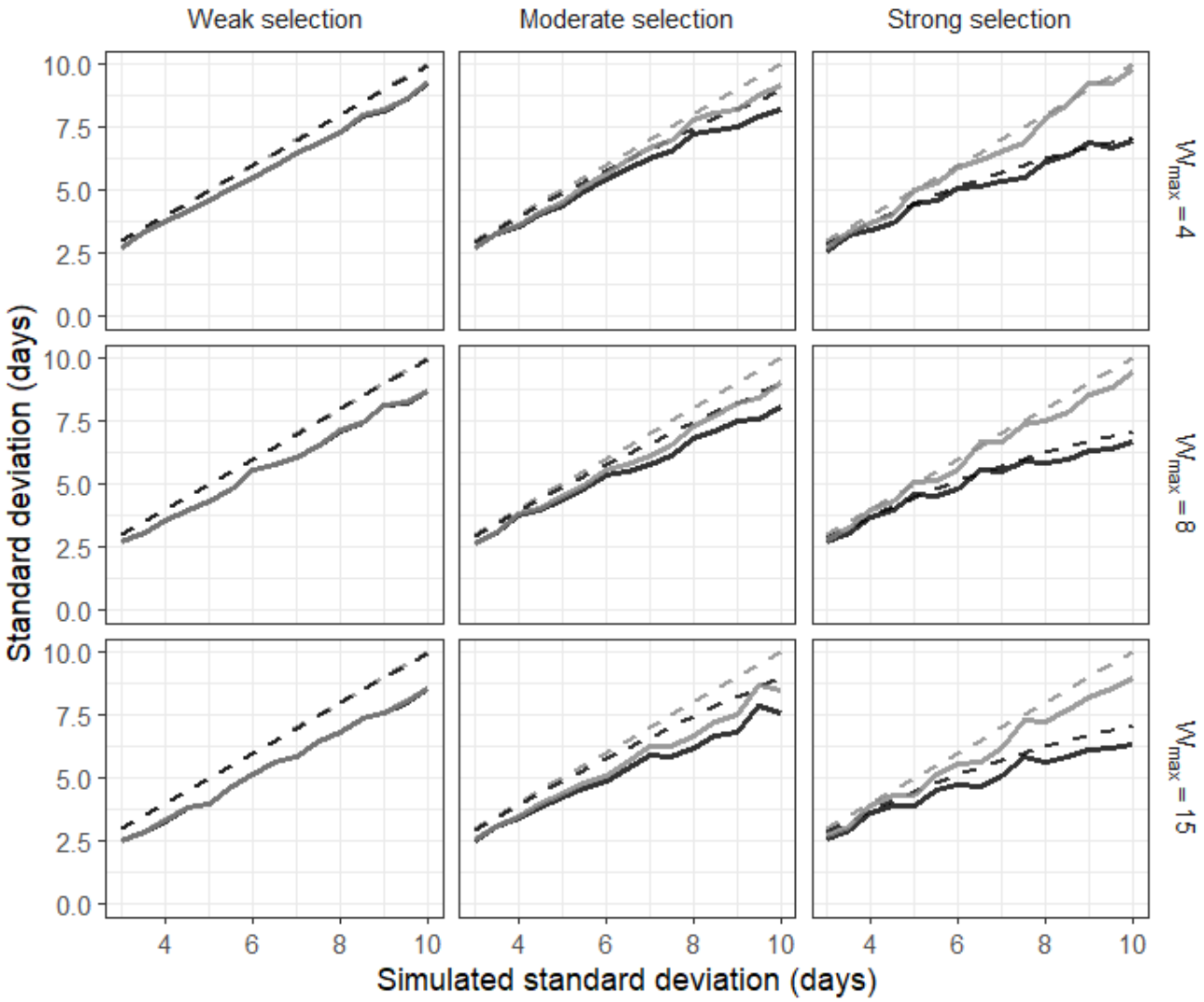


Figure S1: Estimating mean laying dates and fledging time from proportions of juveniles in catches. The mean reproductive phenology, as estimated from the proportions of juveniles in population samples (black lines), or inferred for laying date (dark grey lines), are shown

across a range of the true simulated SD in laying dates, for three strengths of selection (columns, for weak selection: $\omega = 100$; moderate selection: $\omega = 20$; strong selection: $\omega = 10$), and different values of maximum reproductive success (rows). Dashed lines represent the known (simulated) values for each parameter, while solid lines represent the estimated ones. The light grey solid line represents the mean laying date inferred when ignoring the influence of selection, simply subtracting the time to fledge T_f from the mean fledging date.

639 Figure S2: Variance

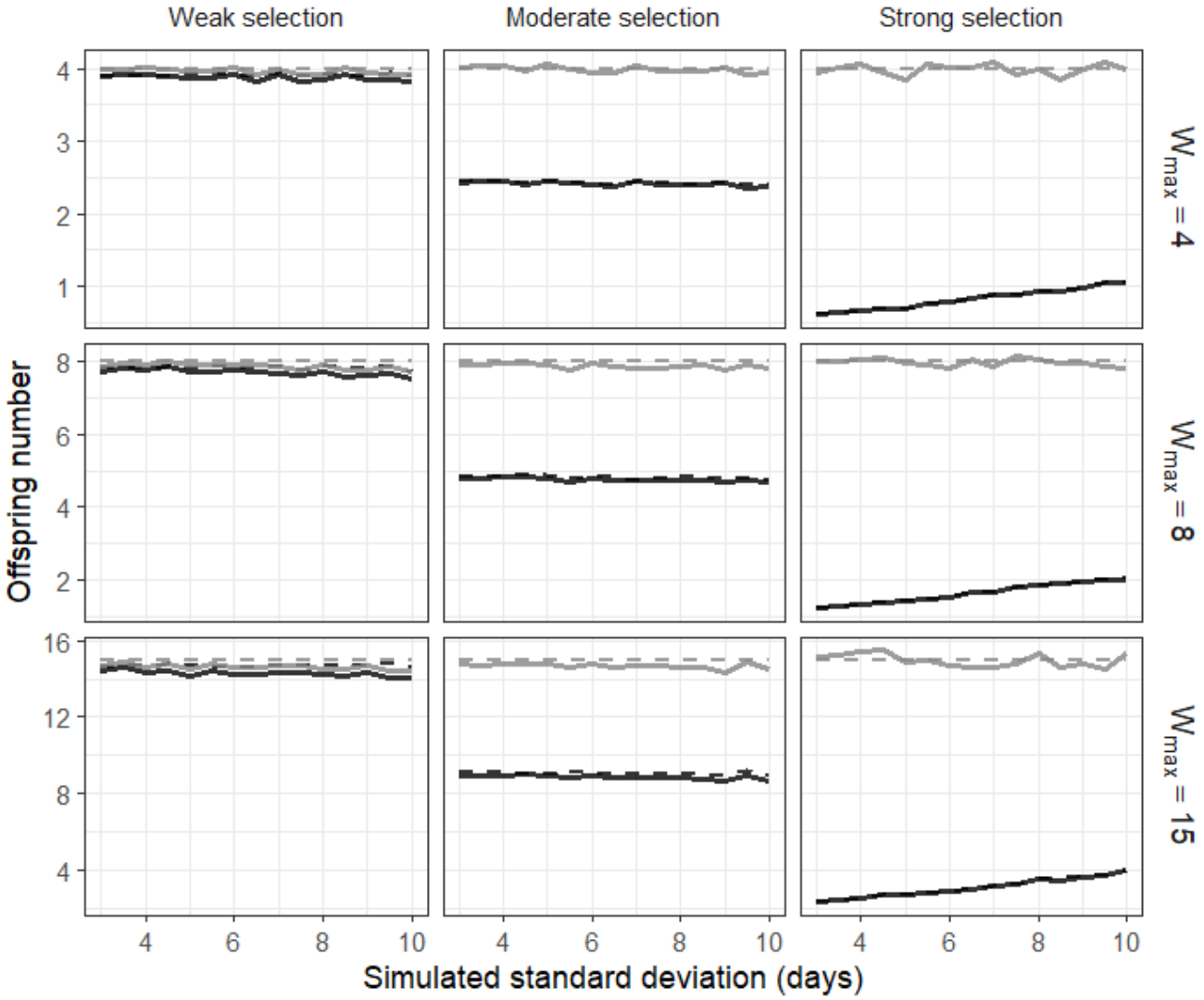


640

641 **Figure S2: Estimating standard deviation in laying dates from proportions of juveniles in**
642 **catches.** The standard deviation of reproductive phenology as estimated from the proportions
643 of juveniles in population samples (black lines), or inferred for laying date (grey lines), are
644 shown across a range of the true simulated SD in laying dates, for three strengths of selection
645 (columns, for weak selection: $\omega = 100$; moderate selection: $\omega = 20$; strong selection: $\omega =$
646 10). Dashed lines represent the known (simulated) values for each parameter, while solid lines
647 represent the estimated ones.

648

649 Figure S3: Realized reproductive success



650

651 **Figure S3: Estimating reproductive success from proportions of juveniles in catches**. The
652 mean the reproductive success as estimated from the proportions of juveniles in population
653 samples (black lines), and the maximum fitness of parents with the optimal breeding date (dark
654 grey lines), are shown across a range of the true simulated SD in laying dates, for three strengths
655 of selection (columns, for weak selection: $\omega = 100$; moderate selection: $\omega = 20$; strong
656 selection: $\omega = 10$). Dashed lines represent the known (simulated) values for each parameter,
657 while solid lines represent the estimated ones.

Appendix 2: Comparing sampling schemes

Figure S4: Between-site variation

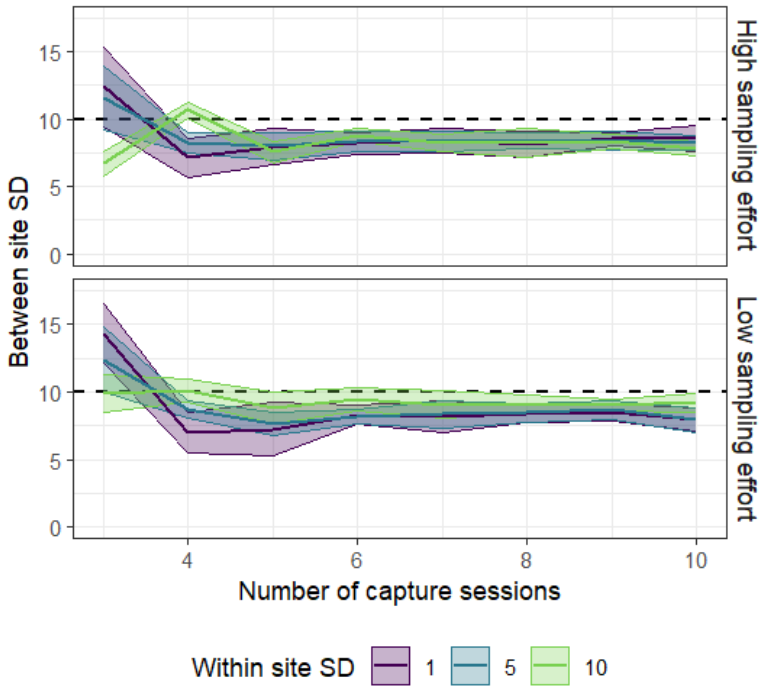


Figure S4: Inferring between-site variation in fledging time. The between-site standard deviation of breeding time, as estimated by the random effect on the midpoint (following eqs. 20-21), is shown across a range of number of capture sessions, for two values sampling effort (number of capturable breeding pairs = 15 for "low sampling effort", upper panel, number of capturable breeding pairs = 200 for "high sampling effort", lower panel) and three values of within site SD (colors). Selection is set to be weak ($\omega = 100$). Dashed line represents the known (simulated) value of between-site variation in fledging time, solid lines represent the estimated ones.

Appendix 3: Model specifications (multi-site)

For each site and for each sampling session t, the number of sampled juveniles $N_{juveniles_{site,t}}$ is assumed to follow a binomial distribution defined by the total number of sampled individuals per site and per session $N_{total_{site,t}}$ and the probability for a sampled individual to be a juvenile $p_{site,t}$ (eq. 1)

$$N_{juveniles_{site,t}} \sim Binomial(N_{total_{site,t}}, p_{site,t}) \quad (1)$$

Where $p_{site,t}$ follows a three-parameters sigmoid function (eq. 2).

$$p_{site,t} = \frac{p_\infty}{1 + \exp\left(\frac{tm_{site} - t}{d}\right)} \quad (2)$$

p_∞ being the asymptote parameter, tm the time at the inflection point and d the slope parameter.

The parameters p_∞ and d are shared between sites and both follow normal distributions (eq. 3-4)

$$p_\infty \sim Normal(\mu_{p_\infty,site}, \sigma_{p_\infty,site}) \quad (3)$$

$$d \sim Normal(\mu_{\mathrm{d},site}, \sigma_{\mathrm{d},site}) \quad (4)$$

The parameter tm varies between sites (eq. 3), and can be considered as the sum of a grand mean (shared among sites) α and a site-specific $\mu_{0,site}$ deviation from this mean (eq. 4).

$$tm_{site} \sim Normal\left(\mu_{tm,site}, \sigma_{tm,site}\right) \quad (5)$$

$$\mu_{tm,site} = \alpha + \mu_{0,site} \quad (6)$$

with $\mu_{0,site}$ being the random site effect (eq. 5), with distribution

$$\mu_{0,site} \sim Normal(0, \sigma_{rd,site}) \quad (7)$$

686 **Priors:**

$$\mu_{p_\infty,site} \sim Normal(0.8,10) \quad (8)$$

$$\mu_{\mathrm{d},site} \sim Normal(2,10) \quad (9)$$

$$\alpha \sim Normal(130,10) \quad (10)$$

$$\sigma_{tm,site} \sim t(0,200) \quad (11)$$

$$\sigma_{p_\infty,site} \sim t(0,100) \quad (12)$$

$$\sigma_{d,site} \sim t(0,10) \quad (13)$$

$$\sigma_{rd,site} \sim Normal(0,60) \quad (14)$$

687