Timing isn't everything: impacts of maximum abundance and duration of a seasonal resource on consumer fitness

Macphie, K. H.[1]*, Pick, J.L. [1], Pearce-Higgins, J.W. [2,3,4], Hadfield, J. D. [1], Phillimore, A.B.[1]

[1] Institute for Ecology and Evolution, The University of Edinburgh, Edinburgh, UK

[2] British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU, UK

[3] Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK

[4] School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

* Corresponding author email: Kirsty.Macphie@ed.ac.uk

Abstract

Phenological advances are among the most apparent biotic responses to warmer springs in mid- to high-latitude regions, with evidence that consumers are advancing less than the resources they rely on. Here, we extend the match/mismatch hypothesis to predict how the mean timing, maximum abundance and duration of the resource phenological distribution impacts on consumer fitness. Using data from 44 sites and 11 years on the phenological distributions of spring arboreal caterpillars and blue tits (*Cyanistes caeruleus*), we find that (1) an earlier caterpillar mean timing is associated with an earlier optimum hatch date, (2) a higher caterpillar abundance peak is associated with higher maximum breeding success and (3) a wider caterpillar distribution is associated with weaker stabilising selection around the optimal hatch date, though the evidence for this third effect is weaker. We then demonstrate that variation in caterpillar distribution height has pronounced effects on blue tit mean breeding success and on the strength of selection for earlier hatch dates. Our analytical approach provides a general framework for examining how phenological distributions of interacting species translate into impacts on mean fitness and selection, key parameters for predicting the fates of populations in a changing climate.

Introduction

Phenology, the timing of seasonal life-history events, defines an organism's interactions with conspecifics, resources, predators and parasites (Macphie & Phillimore, 2024). Over recent decades there have been widespread reports of phenological shifts, particularly through advances of spring events in response to warming (Parmesan & Yohe, 2003), with the extent of advance varying across taxa and regions (Cohen et al., 2018; Roslin et al., 2021). Within food webs there is evidence that the phenology of secondary consumers has advanced less quickly and is less responsive to changes in temperature than the primary producers they feed on (Thackeray et al., 2010, 2016). Continued temperature rises may therefore see secondary consumers lagging further behind key resources (Samplonius et al., 2020; Thackeray et al., 2016; but see Kharouba et al., 2018) leading to concerns that the increasing resource-consumer asynchrony (difference between the mean timing of the resource and the consumer) may impact negatively on consumer fitness (Cushing, 1990; Durant et al., 2005), with knock-on effects on food webs and ecosystem function (IPCC, 2021).

In the context of food webs, phenological research has largely focused on establishing the effects of resource-consumer asynchrony on consumer fitness. Work in this area has generally been framed in the context of the match/mismatch hypothesis (MMH) (Cushing, 1969, 1982, 1990), which was originally proposed to explain how greater asynchrony between the hatching of fish larvae and their zooplankton prey translates into lower availability of resources for the fish and negative consequences for fish growth and recruitment. Since this early work, researchers working on a wide range of systems have gone on to test key predictions that can be made based on the MMH. For instance, several studies find support for the prediction that consumer fitness declines as resource-consumer asynchrony increases in either direction (i.e. where the consumer is too early or too late); typically evidencing this by estimation of a concave quadratic

relationship between fitness and asynchrony (Reed et al., 2013; Youngflesh et al., 2023). Other studies find support for the prediction that in years when the resource is earlier, the asynchrony between resource and consumer is greater, which in turn translates into stronger selection for earlier consumer phenology (Charmantier et al., 2008).

Whilst previous work has focussed largely on asynchrony in the *mean* timing, the resource is distributed through time, typically increasing to a maximum and then decreasing. By holding the mean timing of the resource and consumer constant, it is easy to envisage that variation in the maximum abundance and duration of the resource phenological distribution will affect the amount of resource available to a consumer on a particular date (Durant et al., 2005; Macphie, 2023; Visser & Gienapp, 2019). Rather few studies have considered sensitivity of consumer fitness to aspects of resource availability that go beyond resource-consumer asynchrony alone. Those that have, consider aspects of resource availability, either in terms of the maximum abundance or biomass of a resource (Weir et al., 2026) or in terms of the availability of a resource to particular individual given its phenology (Ramakers et al., 2019). Resource abundance has been shown to play an important role in predicting annual variation in reproductive success of cod (*Gadus morhua*), puffin (*Fratercula arctica*) and Soay sheep (*Ovies aries* L) (Durant et al. 2005), tree swallows (*Tachycineta bicolor*) (Probst et al., 2026), blue tits (*Cyanistes caeruleus*), great tits (*Parus major*), and pied flycatchers (*Ficedula hypoleuca*) (Visser et al., 2015; Weir et al., 2026). However, another study on great tits found that consideration of caterpillar resource availability relative to demand did not improve the ability to predict great tit recruitment or selection as compared with models based on resource-consumer asynchrony alone (Ramakers et al., 2019).

While theoretical work has extended the MMH and demonstrated that the duration of the resource phenological distribution can impact on consumer fitness (Johansson et al., 2015), the

field lacks an analytical framework to test whether the full phenological distribution of a resource (mean timing, maximum abundance and duration; Figure 1) impacts on consumer fitness. Estimating these effects is key to more completely capturing the effects of climate change on resource and consumer phenology and how this in turn translates into shifts in a consumer population's mean fitness and the strength of directional selection on phenology, quantities that define demographic and evolutionary responses (Chevin et al., 2010; Lande, 1976; Lande & Arnold, 1983).

We propose that through estimation of three phenological distributions/functions – (i) resource distribution, (ii) consumer fitness function, (iii) consumer timing – it is possible to arrive at an extended match/mismatch hypothesis (EMMH) that predicts multiple pathways through which resource availability impacts on consumer fitness (Fig 1). First, we have the phenological abundance (or biomass) distribution of the resource, parameterised in terms of mean timing, height (maximum abundance) and width (related to duration, see Fig 1A for details; (Macphie et al., 2023). Second, we have a function that relates the timing of a life-history event in the consumer to its fitness (de Villemereuil et al., 2020; Macphie, 2023) in a way that may depend on the resource phenological distribution (Fig 1A; 26); this function is defined by an optima (θ), maximum fitness (M) and width (ω, Fig 1 A). Where resource availability has a causal effect on consumer fitness, we make three predictions: (*Prediction 1*) mean resource timing will determine the consumer θ, (*Prediction 2*) resource height will determine M and (*Prediction 3*) resource width will determine ω (Fig 1A,B). For the relationship between resource height and M we predict a non-linear saturating function as the fitness benefits of a doubling in resource availability may be much greater when resources are scarce than when they are super-abundant.

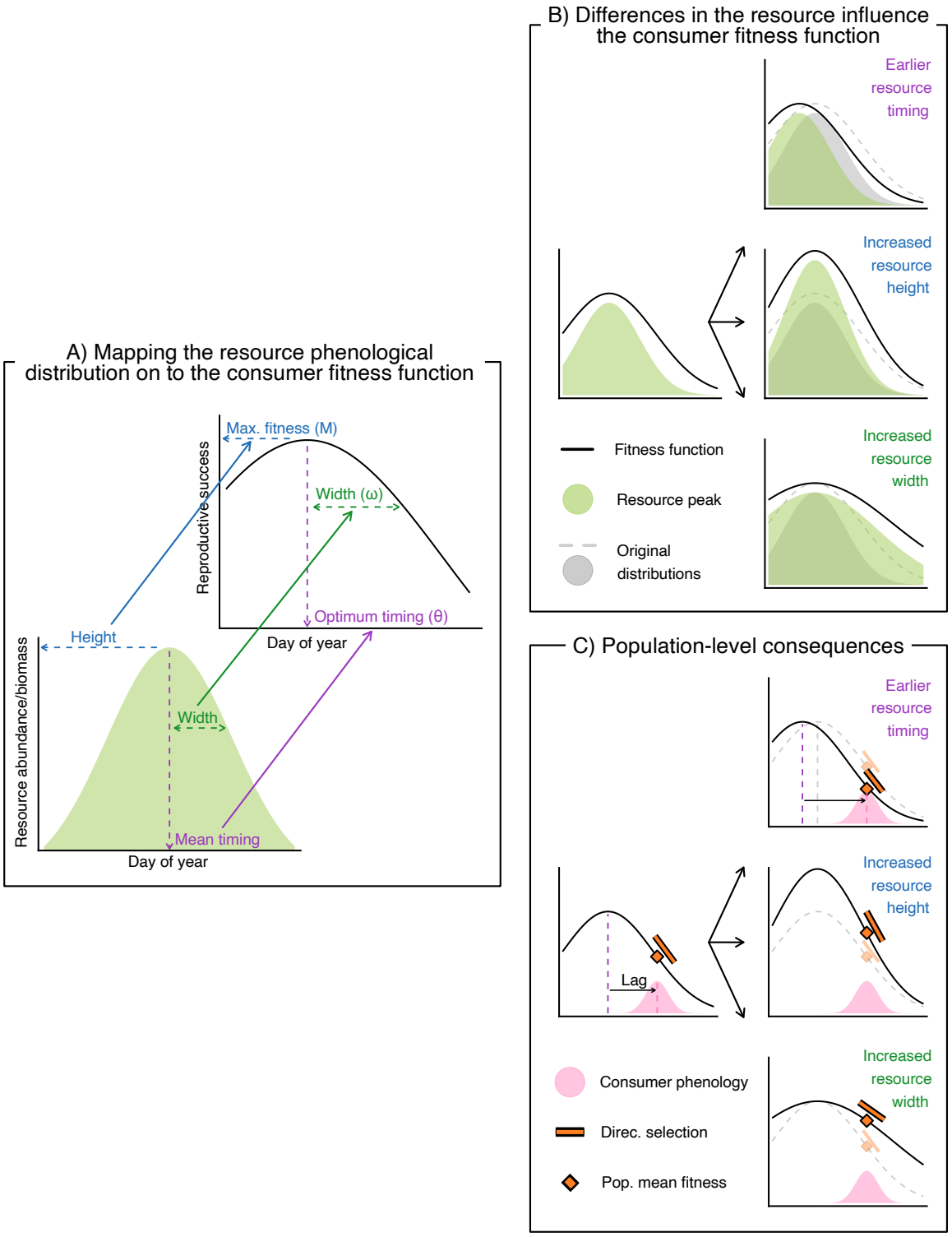


Figure 1. The extended match-mismatch hypothesis (EMMH): A framework for modelling how the phenological distributions of a consumer and resource translate into fitness consequences for consumer populations. A. Predictions of how parameters of the resource

phenological distribution affect parameters of the consumer phenological fitness function. Solid arrows represent causal effects (predicted to be positive). At a fixed height, the duration over which the resource is above a given threshold will be determined by the width. As height increases, the duration above this threshold will also increase. B. Predictions of how changes in resource mean timing, height and width impact the consumer phenological fitness function. C. Illustrates how consideration of the consumer phenological fitness function and phenological distribution together inform lag (difference in timing between θ and mean consumer phenology), mean population fitness (the average of the fitness function weighted across the consumer phenological distribution) and the strength of directional selection on consumer phenology (the expected tangent of the fitness function across consumer phenology).

The third and final distribution is the phenological distribution of the consumer; the distribution of dates on which the life-history event occurs. This distribution can be used to derive the population-level consequences of the estimated fitness function (de Villemereuil et al., 2020) Fig 1C). The difference in timing between θ and the mean timing of the consumer quantifies the lag (referred to as 'mistiming' in (Visser & Gienapp, 2019)). As lag increases, the mean fitness of the population will decline and the strength of directional selection will increase (Fig 1C). At a given lag, increasing M or ω leads to an increase in mean consumer fitness, whilst the strength of directional selection decreases with ω and has the potential to increase with M (Fig 1C). Therefore, considering these three phenological distributions allows examination of the pathways whereby changes in the resource and consumer phenology translate into impacts on lag, components of mean fitness and selection. Quantification of impacts of mismatch on mean fitness are key to understanding short-term demographic implications, whereas impacts on selection are crucial for understanding evolutionary responses (Chevin et al., 2010; Lande, 1976; Lande & Arnold, 1983).

A well-studied model system for understanding how trophic mismatch impacts on consumer fitness is the temperate deciduous woodland food chain. A spring pulse in the abundance and biomass of caterpillars are heavily relied on by insectivorous birds whilst provisioning young in the nest. In this system the timing of the spring peak caterpillar abundance varies across time and space (Macphie et al., 2025), arising earlier in years and locations where spring is warmer (Macphie et al., 2023; Visser et al., 2006). Similarly, the nesting phenology (egg-laying, hatch dates and peak nestling demand) of insectivorous passerines, like blue tits, great tits and pied flycatchers, is also advanced when spring is warmer (Phillimore et al., 2016). As a result, the timing of the peak demand of nestlings partially tracks the peak timing of caterpillar abundance and biomass across time and space (Burgess et al., 2018; Charmantier et al., 2008). The peak energetic demands of tit chicks are approximately 10 days after hatching (Perrins, 1991), with nests where peak demand is substantially earlier or later than the caterpillar distribution associated with a reduction in fledging success, recruitment of offspring to the population and adult survival (Charmantier et al., 2008; Reed et al., 2013). The slower phenological advance shown by some of these bird species relative to their caterpillar resource (Both et al., 2009) has exposed them to increasing trophic asynchrony. As discussed above, most work testing the MMH has focused on the effects of mean timing, overlooking causal predictions that can be made for effects of the height and width of the full caterpillar distribution (Fig 1A). Not only can testing these predictions increase our confidence that resource availability is having a causal effect on consumer fitness, but they also represent additional indirect pathways for effects of spring temperature on consumer fitness via the resource distribution (Macphie et al., 2023).

Here, we test our three EMMH predictions using data on the phenological distribution of caterpillars and the hatching phenology and fledging success of blue tits across 1,366 breeding attempts, 11 years and 44 woodland sites across Scotland (Macphie et al., 2023; Shutt et al., 2018). We jointly estimated the phenological distribution of caterpillars (hereafter, caterpillar

distribution) and the phenological fitness function for blue tit hatch date, estimating the effects of parameters that govern the caterpillar distribution on parameters that govern the hatch date fitness function (Fig 1A). Extending the MMH to incorporate the full resource distribution holds potential to identify additional pathways via which an ephemeral resource can impact consumer fitness, increasing our confidence in the causality of identified effects, and improve the accuracy of our estimates of population mean fitness and the strength of selection on reproductive phenology in a rapidly changing world.

Methods

Study System

We collected data over 11 years, from 2014 to 2024, at 44 deciduous woodland sites along a 220km transect between Edinburgh (55°980 N, 3°400 W) and Dornoch (57°890 N, 4°080 W) in Scotland. The sites span an elevational range of 10-433m above sea level and vary in woodland tree composition and temperature (Shutt et al., 2018). From 2014-2016 between 30 and 37 sites were monitored, the majority of which had six nest boxes (26mm hole Schwegler 1B) to target breeding blue tits. In 2017 we increased coverage to 44 sites, with most sites having eight nest boxes, though seven sites had six boxes and one had four. From 2020-2024 onwards the number of sites intensively monitored varied across years, with 22,37,43,23,22 sites monitored respectively. Each site was visited every second day from 1st April each year (except 2020 when, due to covid-19, monitoring began on 7th May), until late June.

Arboreal caterpillar abundance

We used branch beating to sample the abundance of arboreal caterpillars present from late spring into early summer (Shutt et al., 2019). We extended a clear rubble sack (76mm x 51 mm)

over a marked branch, and hit the branch firmly 30 times to dislodged caterpillars, recording the abundance of caterpillars ≥1mm diameter. An average of 14 trees were surveyed at each site in each year from 2017-24 (five trees per site prior to that) with half of the trees at each site sampled on alternating visits, meaning that every tree was beaten every four days. The four-day interval allowed caterpillars to recolonise between sampling events. The same branches were sampled throughout each season and across years unless they were broken or dead. We recorded the leafing phenology of the marked trees at each site, and branch beating started once 45% of trees across all sites had unfurled the first leaf, with sampling continuing until the end of the field season in mid/late June (see Bialy et al., 2026), for more details on annual sampling periods).

Blue tit breeding phenology and success

At each site, we checked nest boxes every second day to monitor blue tit breeding phenology, recording the progression of clutch size and the onset of incubation. Once incubation had begun, the nest was undisturbed for 10 days, after which we resumed checks on alternate days to determine the hatch date within a two-day window of uncertainty. The ordinal date on which hatched chicks were first observed in the nest was used as the hatch date in our analyses. 265 nest records from 2017-19 were removed due to involvement in a cross-fostering experiment.

We fitted nestlings with a unique ring, under licence from the British Trust for Ornithology, six days after hatching was first observed. Nests were visited again 12 days after the observed hatch date and we recorded which individuals were alive. From 18 days after the observed hatch date, we monitored nests every second day for fledging until all nestlings had left the nest. Once fledging was complete, we searched nest contents to identify dead nestlings, and from this identified the number of young that had fledged.

To identify the female bird breeding at each nest box, we caught adult birds on the nest ≥10 days after hatching. For one nest, two females were caught, so for this nest we assigned a unique female identity. Female identity is known for 85% of all nests and was treated as unique for the remaining 15%. A total of 1,366 nests were included in our analyses.

Extending the MMH using the Gaussian function

Phenological distribution of caterpillar abundance: To use the caterpillar distribution parameters as predictors of the blue tit phenological fitness function, we first modelled the caterpillar distribution (phenological distribution of abundance) in each site in each year. The Gaussian function is well suited to describing the phenological distribution of caterpillar abundance (C) because it is governed by three parameters which define the mean timing (θ_C), maximum height (M_C) and width (ω_C) of the caterpillar distribution (Macphie et al., 2023):

Eq. 1: $$\mathrm{E}[C(x)] = \mathrm{M}_C exp\left(-\frac{(x-\theta_C)^2}{2{\omega_C}^2}\right)$$

Eq. 1 can be rearranged to model caterpillar abundance on the log scale, where f_C is the log transform:

Eq 2: $$f_C(\mathrm{E}[C(x)]) = f_C M_C - \frac{(x-\theta_C)^2}{2\exp{(ln\omega_C)^2}}$$

We fitted a Poisson model to the number of sampled caterpillars where the expectation was given by Eq 2 and the date (x) was centred to the midpoint of the caterpillar and blue tit hatching data: day 146 (25th or 26th May). The three phenological parameters (θ_C, $f_C M_C$, $ln\omega_C$) were

modelled using a trivariate mixed model with a fixed intercept and random site, year and site by year interaction effects fitted for each phenological parameter. These allow the mean timing, log height and log width of the caterpillar distribution to be estimated for each site in each year. The site effects (s) for the three phenological parameters were treated as normal with estimated covariance matrix, $\boldsymbol{V}_{C:s}$. $\boldsymbol{V}_{C:s}$ can be decomposed into a correlation matrix ($\boldsymbol{R}_C$) and a diagonal matrix of site-effect standard deviations ($\boldsymbol{D}_{C:s}$): $\boldsymbol{V}_{C:s} = \boldsymbol{D}_{C:s}\boldsymbol{R}_C\boldsymbol{D}_{C:s}$. The covariance structure for the year effects (y) and site by year effects (sy) share the same correlation matrix ($\boldsymbol{R}_C$) but have unique standard deviations ($\boldsymbol{D}_{C:y}$ and $\boldsymbol{D}_{C:sy}$). We also included a day-by-site-by-year interaction, the recorder of each sample, unique tree identity and an observation-level effects as random terms for $f_C M_C$ to account for other sources of variation in caterpillar abundance, as in (Macphie et al., 2023).

Phenological fitness function: As 14% of nests fledged zero offspring, but the average number of offspring fledged per nest was 6.30 among those that did not fail, the data were zero-inflated relative to a Poisson distribution. To account for this, we used a hurdle model composition, estimating a fitness function for two components of fledging success: (i) success or failure in fledging at least one individual, and (ii) conditional on success, the number fledged. We modelled the probability of success/failure of a nest (S) as a function of hatch date (h):

Eq. 3: $$f_S(E[S(h)]) = f_S\mathrm{M}_S - \frac{(h-\theta_S)^2}{2exp(ln\omega_S)^2}$$

Where f_S is the logit function. The parameterisation follows that of the Gaussian function in (de Villemereuil et al., 2020) (Eq. 2) where we estimate the optimum timing (θ_S), the maximum probability of success on the logit scale ($f_S\mathrm{M}_S$) and the log width ($ln\omega_S$), again centring hatch date on day 146. As in the caterpillar abundance model, we fitted site, year and site by year

effects as random for all three phenological parameters with a common correlation matrix ($\boldsymbol{R}_S$) but effect-specific standard deviations ($\boldsymbol{D}_{S:s}, \boldsymbol{D}_{S:y}$ and $\boldsymbol{D}_{S:sy}$). The female identity for each nest was also included as a random term for $f_S\mathrm{M}_S$, allowing the fledging success to vary among females.

For successful nests, the variance of the number fledged was 5.95; this is under-dispersed with respect to a Poisson distribution, especially given that we know there are multiple factors driving variance in the number fledged. Therefore, we modelled the number fledged (N), conditional on success, as a function of hatch date, using a zero-truncated generalised Poisson distribution (TGP). The model has the same form as nest success, but with a log link function:

Eq. 4: $$f_N(E[N(h)]) = f_N\mathrm{M}_N - \frac{(h-\theta_N)^2}{2exp(ln\omega_N)^2}$$

The generalised-Poisson distribution is also parameterised by an additional parameter that defines the over- or under-dispersion, where 0 = a standard Poisson distribution (Consul & Jain, 1973). As in the nest success model, we fitted site, year and site by year effects as random for all three phenological parameters. For site effects, their covariance structure was modelled as $\boldsymbol{B}\boldsymbol{V}_{S:s}\boldsymbol{B} + \boldsymbol{D}_{N:s}\boldsymbol{R}_N\boldsymbol{D}_{N:s}$ where $\boldsymbol{B}$ is a diagonal matrix of regression coefficients relating the fledging number site effects (response) on the nest success site effects (predictor) for each phenological parameter. The covariance structure for the year and site by year effects have the same form with the same $\boldsymbol{B}$ and $\boldsymbol{R}_N$. However, effect-specific standard deviations are fitted (ie. $\boldsymbol{D}_{N:y}$ and $\boldsymbol{D}_{N:sy}$). Non-zero regression coefficients allow covariances to exist between effects on fledging number and nest success. For example, the covariance matrix between site effects on fledging number and site effects on nest success is $\boldsymbol{B}\boldsymbol{V}_{S:s}$. Due to convergence issues the

287 regression coefficient associated with the widths was set to zero. The female identity for each
288 nest was also included as a random term for $f_N \mathrm{M}_N$.
289
290 **The caterpillar distribution as a predictor of blue tit fitness:** The predicted caterpillar
291 phenological parameters for each site in each year were used as predictors of the parameters
292 appearing in the blue tit phenological fitness functions (Eqs. 3 and 4). These predictions
293 resulted in three focal slopes: (1) a linear effect of θ_C on θ_S and θ_N which allowed the mean
294 timing of the caterpillar distribution to affect the optimum timing of the blue tit phenological
295 fitness functions, (2) a linear effect of on $f_C \mathrm{M}_C$ on $f_S \mathrm{M}_S$ and $f_N \mathrm{M}_N$which allowed the maximum
296 height of the caterpillar distribution to affect the maximum fledging success and number of
297 fledglings at the optimum time of breeding for blue tits, (3) a linear effect of $ln\omega_C$ on $ln\omega_S$ and
298 $log\omega_N$ which allowed the width of the caterpillar distribution to affect the width of the fitness
299 functions. These slopes were predicted to be positive and constitute *Predictions 1-3* respectively
300 (see above). We also included an effect of $f_C \mathrm{M}_C$on on $ln\omega_S$ and $ln\omega_N$, which allowed the height
301 of the caterpillar distribution to affect the width of the blue tit fitness functions. We included two
302 slopes on the width of the fitness function because the height and width parameters both
303 influence the duration over which the caterpillar distribution falls and these parameters
304 negatively covary (Macphie et al., 2023).
305
306 **The extended match/mismatch hypothesis model:** The *Fitness model* was coded in RStan
307 (Stan Development Team, 2022) within one framework, co-estimating the caterpillar
308 phenological distribution parameters from the raw abundance data and the blue tit phenological
309 fitness function for the Bernoulli and TGP components. This framework allowed any uncertainty
310 in the caterpillar distribution estimates to be incorporated into the model. See Supporting
311 Information (Table S1) for model output summary.

The *Fitness model* was run in RStan using four chains and 24000 iterations thinned by 20 and including a warmup of 2000. We aimed to use relatively uninformative priors for all coefficients. For fixed effects, we used normally distributed priors with a mean of zero and standard deviations of five, 10 or 20, depending on the expected scale of the coefficient - see Table S1. The intercept for the logit-scale maximum fitness of the Bernoulli part of the model had a standard deviation of $\sqrt{\pi^2/3}$, which prevents the prior distribution on the latent-scale pulling estimates towards zero or one. For random terms, we use half-cauchy priors on the standard deviation with a location of zero and scale of 10 for most terms, but a scale of three for the fitness function height and width random terms to aid with model convergence. The prior for the Cholesky factor of the covariance correlation matrices had an eta of 2. The *Fitness model* had 29 divergent transitions after warmup; 0.66% of the 4400 iterations retained. We used posterior predictive simulates to check model fit, see Supporting Information for details.

Lag: To estimate the lag for each population in each year, we first estimated the population mean hatch date for each site-year combination using a linear mixed model (*Hatch date model*). The model was run using brms (Bürkner, 2017) with centred hatch date as the response variable, an intercept and random terms of site, year, a site by year interaction and female ID. We ran four chains of 12000 iterations with a warmup of 1000 iterations and thin of 10 and default priors. From this model, the site, year and site by year effects were combined to produce a posterior distribution of the predicted mean hatch date at each site in each year. The MCMC chain was equal in length (iterations) to that of the *Fitness model*. We subtracted the optimum hatch date for each site-year from the mean hatch date for each site-year across the posterior to produce a posterior distribution of estimated lag.

Mean fitness and selection:

We examined variation in two components of population mean fitness (here referred to as mean fitness) and directional selection from two perspectives: 1) To quantify the contribution that the caterpillar distribution makes to spatiotemporal variation in mean fitness and directional selection, we estimated the proportion of variation in each that is explained by the caterpillar distribution across sites, years and site-years. 2) To understand the effect of caterpillar distribution height and width on mean fitness and directional selection, we generated predictions for hypothetical populations at different amounts of lag. We considered the probability of success and number fledged independently.

1) Estimates of the population mean fitness and directional selection at each site in each year were calculated across the posterior using site-year specific predictions of the caterpillar distribution and fitness functions, from the *Fitness model*, and predicted mean hatch date, from the *Hatch date model*. As we have a maximum of eight nests per site-year, for each iteration, we drew 1,000 hatch dates from a normal distribution with a mean of the site-year estimated mean hatch date and standard deviation (s.d.) of the within site-year s.d., as estimated in the *Hatch date model*. Using all model terms, specific to the relevant site-year where appropriate, we calculated mean fitness for each hatch date on the latent scale. The mean fitness across hatch dates was stored as the mean fitness and each fitness estimate was divided by the mean, relativising fitness. Hatch date was mean centred on zero and fitness was modelled as a quadratic function of centred hatch date in a linear model. The linear hatch date coefficient was stored as the estimate of directional selection (Lande & Arnold, 1983). By performing these steps across every iteration, we obtained posterior distributions for the estimates of population mean fitness and directional selection. We estimated population mean fitness and directional selection twice, first using only the fixed effects in which the caterpillar distribution predicted variation in the fitness function (*caterpillar predictions*), and second using all model terms for

each site in each year including the effects of the caterpillar distribution but also the remaining variation within and between site-years (*complete predictions*). For each iteration, we used linear mixed models to decompose the variance in mean fitness and directional selection among sites, years and the site by year interaction. The total variance among *caterpillar predictions* of fitness and selection were divided by the variance among *complete predictions* to describe the proportion of variance in mean fitness and directional selection explained by the caterpillar distribution. The site, year and residual (site by year) variances were used to consider the spatio-temporal contribution to variance in population mean fitness and directional selection.

2) To examine the consequences of the EMMH on mean fitness, directional selection and their association with lag, we made predictions across a range of caterpillar heights, widths and lags. We focus on caterpillar distribution height and width, as the novel components within the EMMH; the caterpillar mean timing would inform the optimum and therefore is incorporated via the lag. We estimated mean fitness and directional selection for hypothetical populations with lags of zero, seven and 14 days behind the optimum, spanning much of the range observed within our data. We used the 0.1, 0.5 and 0.9 quantiles of the caterpillar distribution height and width among site-year mean estimates (points in Figure 2) and varied height or width independently, whilst holding the other caterpillar distribution metrics at their median values. We estimated mean fitness and directional selection across each iteration and each lag and height or width combination using lags and fitness functions aligned with our hypothetical populations. We sampled 10,000 hatch dates from a normal distribution with a mean of the relevant lag and residual s.d. estimated in the *Hatch date model*. We drew samples of random effects from the variance estimates for each *Fitness model* term for each sampled hatch date. We then followed the same process outlined in (1) to estimate mean fitness and directional selection.

Results

Match/mismatch and the fitness function

Using data from 1,366 blue tit nests and 51,660 samples of caterpillar abundance across 369 site-years, we identified strong support for positive relationships between optimum timing and the maximum fitness and their corresponding parameters describing the caterpillar distribution (EMMH *predictions 1* and *2*) and partial support for a positive effect of width of the resource phenological distribution on the width of the consumer fitness function (EMMH *prediction 3*).

For each day advance in the mean timing of the caterpillar distribution, the optimum hatch date shifted earlier by 0.64 (posterior mean, with 95% credible intervals: 0.22 – 1.07) days for the probability of success (Figure 2b) and 0.82 (0.49 - 1.14) days for the number fledged (*prediction 1*; Figure 2f). In both cases the estimated slope was shallower than 1:1, though credible intervals overlap a slope of 1. Across site-years, caterpillar mean timing explained 71.03% (11.40 – 98.96) and 58.65% (20.07 – 92.65) of the variance in optimum timing for the probability of success and the number fledged, respectively. For site-years with an earlier caterpillar distribution, e.g. day 140 (20th May in a non-leap year), we found that it was optimal to hatch 11.8 days (4.6 – 22.1) earlier than the caterpillar peak to maximise the probability of successfully fledging some young, whilst the number of young fledged was maximised when hatch dates are 15.0 days (9.5 – 22.4) earlier than the peak. In comparison, for site-years with a late caterpillar distribution, e.g. day 170 (19th June in a non-leap year), the hatching date fitness optimum was 22.5 days (14.5 – 34.7) and 20.5 days (13.3 – 30.6) before the peak for the probability of success and number fledged, respectively.

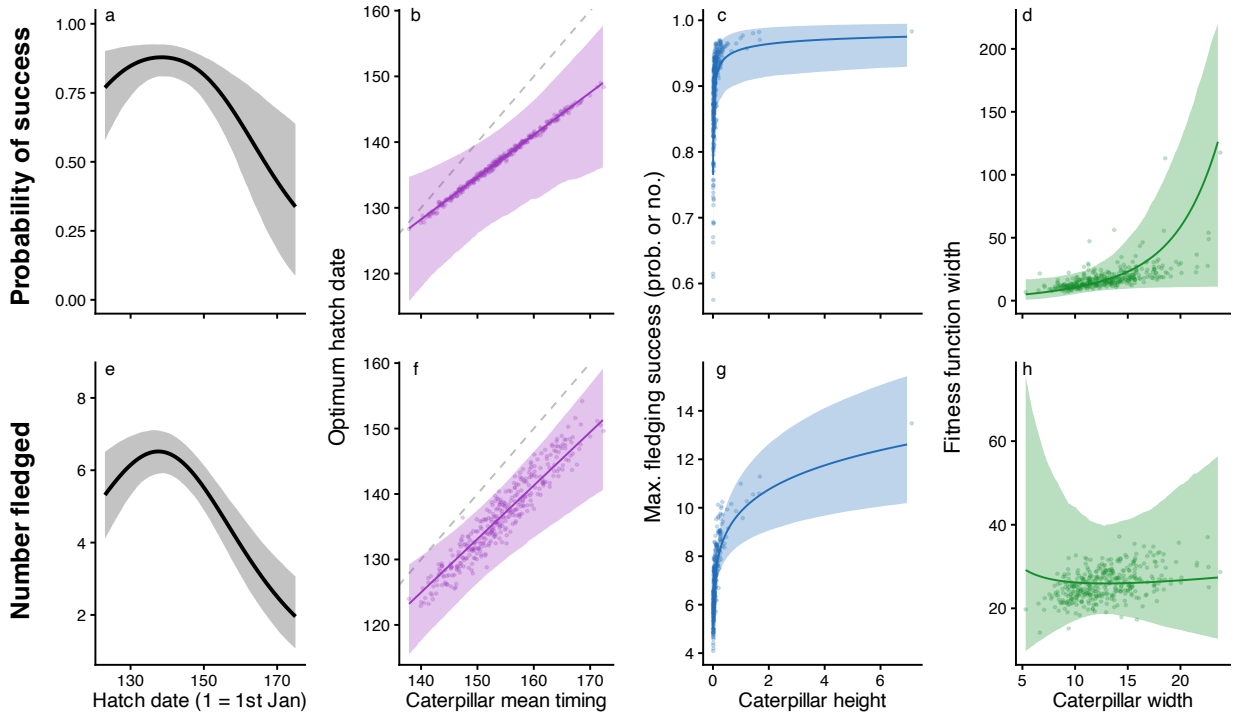


Figure 2. The mean hatch-date phenological fitness function (a,e) and relationship between parameters for caterpillars and blue tits (b-d, f-h) estimated across site-years, for probability of success (a-d) and number fledged (e-h). In all panels the solid line corresponds to the mean of the posterior on the data scale and shaded areas correspond to 95% credible intervals. In b-d and f-h the points represent site-year mean predictions from the model. The grey-dashed line in b and f corresponds to the timing expected based on the maximum resource demand of chicks (caterpillar mean timing - 10 days, when the resource demand of chicks is highest; (Perrins, 1991). See Figure S2 for c and g with a cropped x-axis, showing trend through narrower range of most common caterpillar height estimates.

The effect of the caterpillar distribution height on the maximum of the hatch date fitness function was positive (*prediction 2*) and non-linear on the data-scale, saturating at higher values, for both the probability of success (0.70; 0.24 – 1.27; Figure 2c) and the number fledged (0.13; 0.09 –

0.17; Figure 2g). For number fledged, our slope estimate is substantially < 1, which we interpret as strong evidence for diminishing returns of increasing caterpillar availability. For probability of success, diminishing returns are imposed by the link function for a positive non-zero coefficient. Across site-years the height of the caterpillar distribution explained 33.02% (4.07 – 72.31) of the variance in maximum fitness at the optimum for the probability of success, with predictions of the maximum probability of success ranging from 0.77 (0.63 – 0.87) to 0.97 (0.93 – 0.99) from the lowest to highest caterpillar peaks. Similarly, across site-years the height of the caterpillar distribution explained 57.90% (28.29 – 83.03) of the variance in maximum fitness at the optimum for the number fledged, with predictions ranging from a maximum of 5.00 (4.36 – 5.71) to 12.61 (10.19 – 15.43) offspring fledged from the lowest to highest caterpillar peaks.

The effect of the width of the caterpillar distribution on the width of the hatch date fitness function (*prediction 3*) was less pronounced. For the probability of success, 0.97 of the posterior distribution of the slope was positive (Figure 2d), consistent with a wider fitness function when the caterpillar distribution is wider. Across site-years the width of the caterpillar distribution explained 51.23% (0.69 – 95.88) of the variance in the width of the fitness function for the probability of success. We found no effect of the caterpillar distribution width on the width of the phenological fitness function for the number fledged (0.01; -1.05 – 1.04; Figure 2h). To control for the height of the caterpillar peak whilst estimating the effect of the width of the caterpillar peak on the width of the fitness function, we included a slope allowing for an association between caterpillar distribution height and the width of each fitness function (Fig S1). We found no clear relationship between the caterpillar distribution height and fitness function width for either the probability of survival (0.21; -0.08 – 0.60) or the number fledged (-0.09; -0.29 – 0.11).

Lag, mean fitness and selection

Next, we quantified among-site, -year and -site-year variation in lag, two *components* of population mean fitness (mean probability of success and mean number fledged – hereafter referred to as 'mean fitness') and selection gradients. To do so, we used a linear mixed model of hatch date to estimate the mean hatch date at each site in each year. Estimates of site-year mean hatch dates and optimum hatch dates were used to calculate predictions of lag (Fig 1c). Estimates of site-year mean hatch dates and the residual variance in hatch dates within a site year were used to simulate annual population-level estimates of the probability and success and number fledged, from which mean fitness and selection gradients were calculated.

For the average site-year, the lag was positive for both probability of success (lag = 7.30 days) and number fledged (lag = 8.00 days), meaning that mean hatch dates tended to be later than the optimum. For the probability of success, site-year level lag varied between -1.35 and 16.53 days (Figure 3a) and for the number fledged, site-year level lag varied between -5.26 and 20.82 days (Figure 3b). We found that for number fledged, the site-year mean lag was elevated when the caterpillar distribution fell earlier (Figure 3f), whereas this relationship is less apparent for the probability of success (Figure 3e). When considering lag between each individual nesting attempt and the optimum timing within a site-year, we observe a greater spread of values from -9.99 to 30.79 and -10.37 to 32.21 days for probability of success and number fledged, respectively (Figure 3c,d).

The average mean fitness among site-years for the probability of success and the number fledged were 0.86 (0.79 – 0.91) and 6.07 (5.46 – 6.64), respectively. Across site-years the caterpillar distribution explained 32.20% (7.57 – 71.06) of the total variance in the probability of success and 44.34% (25.23 – 67.74) of the total variance for the number fledged (Figure 4, S3).

For both fitness components we find that site, year and site-year effects contribute to (i) the total variance among site-years and (ii) the partial variance that is attributable to effects of the caterpillar phenological distribution (Figure 4).

The average selection gradient for hatch date across site-years was -0.006 day^{-1} (-0.014 – -0.001) and -0.015 day^{-1} (-0.025 – -0.006) for probability of nesting success and number fledged, respectively. Across site-years, the caterpillar distribution explained 36.67% (1.03 – 139.24) of the variance in directional selection for the probability of success and 37.76% (8.22 – 84.98) of the variance in number fledged (Figure 4, S3). For both fitness components, we find that site, year and site-year effects contribute to the (i) total variance in mean fitness among site-years and (ii) the partial variance that is attributable to effects of the caterpillar phenological distribution (Figure 4).

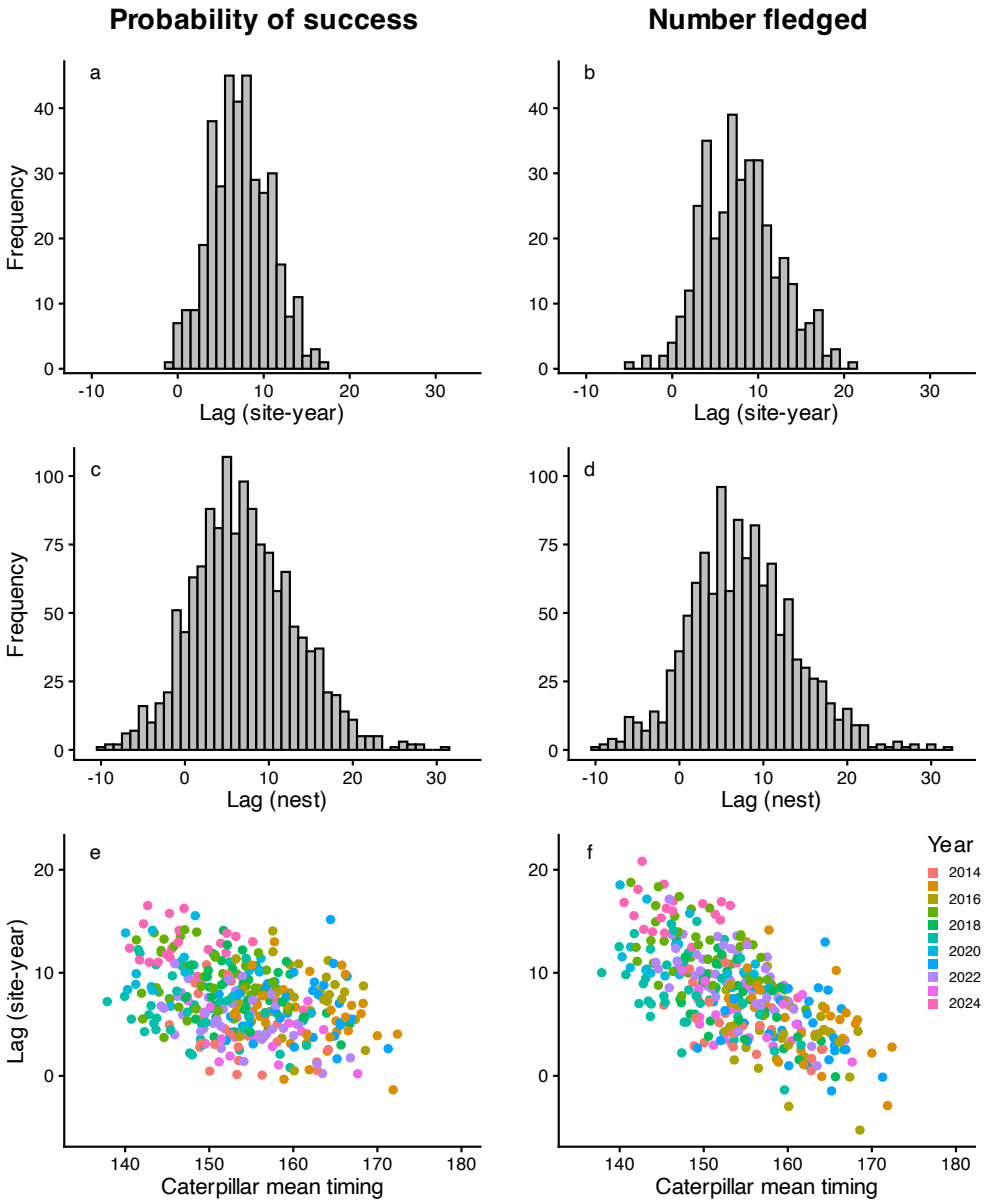
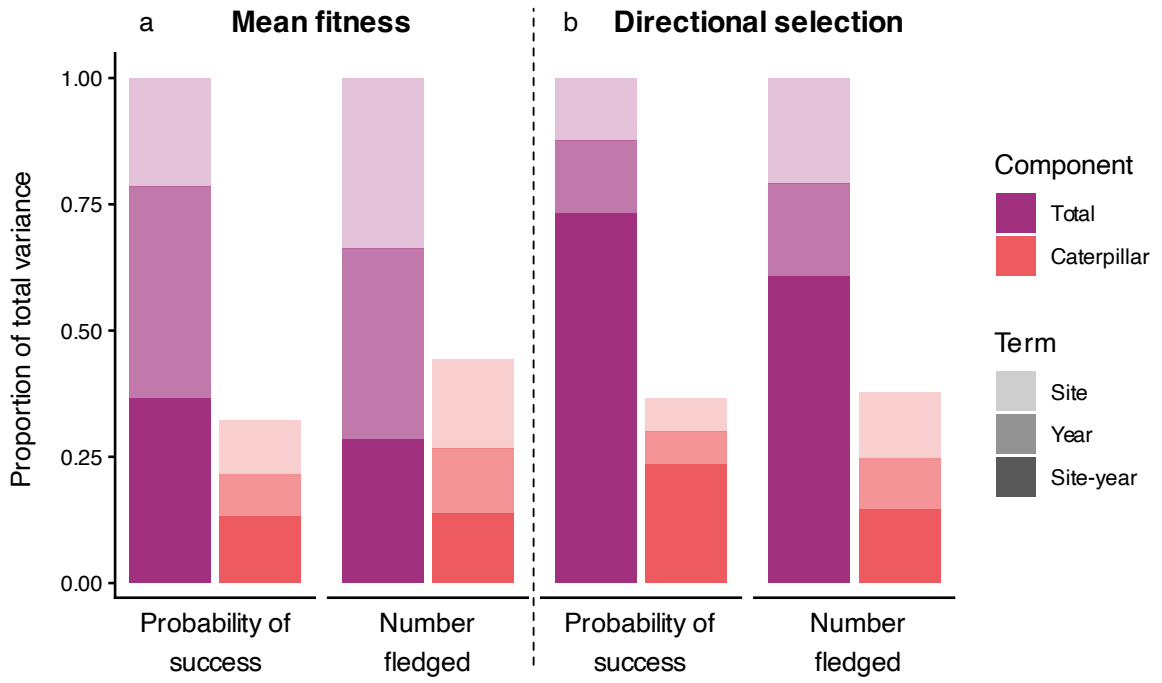


Figure 3. Histograms of the lag (hatch date – optimum) estimate for each site-year (a,b) and nest (c,d) and the relations of lag with caterpillar mean timing across site-years (e,f) for probability of nesting success (a,c,e) and number fledged (b,d,f). Colours of datapoints in e and f correspond to year.



498

499 **Figure 4: Spatiotemporal decomposition of the variance in a) population mean fitness**
500 **and b) directional selection from (i) total site-year variance estimates and (ii) partial site-**
501 **year estimates attributable to variation in the caterpillar distribution.** We present the
502 posterior mean estimates. The summed height of the red bars represent the proportion of the
503 total variance in mean fitness and directional selection that can be explained by the caterpillar
504 phenological distribution.

505

506 Under the EMMH framing we expect the effects of the lag on mean fitness and directional
507 selection to be sensitive to the height and width of the caterpillar distribution (Fig 1c); with
508 caterpillar mean timing accounted for within lag via the optimum. To assess this, we used our
509 fitness function model to predict mean fitness and directional selection at different lags and
510 caterpillar distribution heights and widths for hypothetical populations. We did so across values
511 identified within predictions from our data (Figure 2, 3), at lags of zero, seven and 14 days and
512 at the 0.1, 0.5 and 0.9 quantiles of the caterpillar distribution height and width estimates across
513 site-years.

In general, we found that both fitness components decline as lag increases at all levels of caterpillar height and width (Figure 5a,b). Site-year level population mean fitness increased with the height of the caterpillar phenological distribution at all levels of lag (Figure 5a), driven by the positive effects of the caterpillar height on blue tit maximum fitness (Figure 2). For the probability of fledging success, the height of the caterpillar distribution had the most pronounced effect on population mean fitness at higher lags, whilst for the number fledged, the effect was most pronounced where lag was zero (Figure 5a). The effect of lag on mean fitness differed with caterpillar distribution height for both fitness components; the rate of decline in mean fitness with lag was more pronounced at lower caterpillar distribution heights for the probability of success and at higher caterpillar distribution heights for the number fledged (Figure 5a).

Differences in mean fitness attributable to the caterpillar distribution width were only apparent at higher lags for the probability of success, however the difference in mean fitness estimates obtained for the 0.1 and 0.9 caterpillar width quantiles overlapped zero (Figure 5b). There was, however, a difference in the rate of decline in mean fitness with lag for the probability of success, with a more pronounced effect of lag at lower caterpillar widths (Figure 5b).

In general, we find that directional selection for earlier hatch dates strengthens as lag increases, for both components of fledging success at all levels of caterpillar height and width (Figure 5ab). While effects of the height and width of the caterpillar phenological distribution on the strength of directional selection were more apparent for the probability of success than the number fledged (Figure 5ab), in no case was the effect significant. The effect of lag on the strength of directional selection differed with caterpillar distribution height and width for the probability of success, with a more pronounced increase in the strength of directional selection with lag at lower caterpillar distribution heights and widths (Figure 5ab).

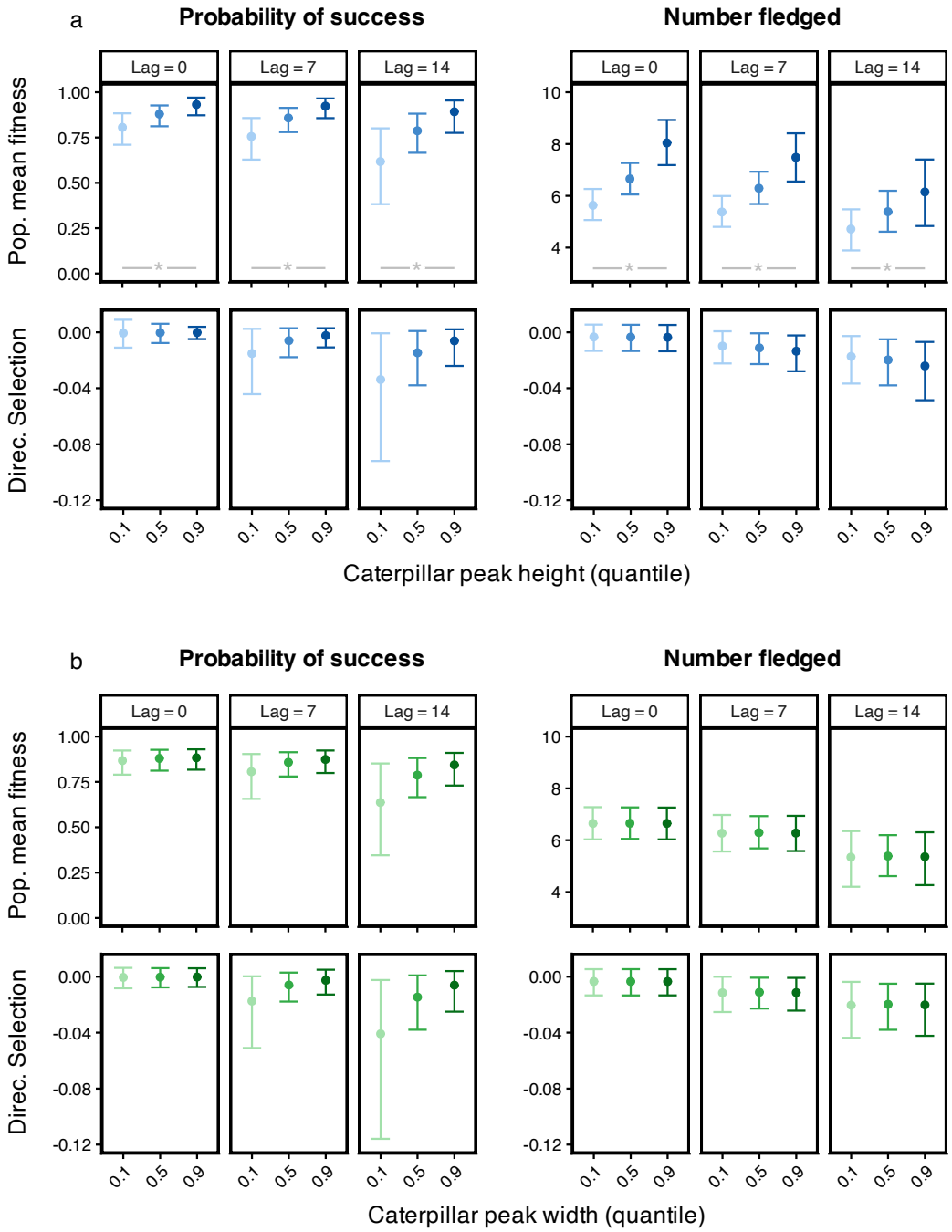


Figure 5. Predictions of mean fitness and directional selection with varying lag and caterpillar height (a) and width (b) for two components of blue tit fledging success (mean and 95% credible intervals). We took the 0.1, 0.5 and 0.9 probability quantiles from the range of site-year lag predictions for our populations, ensuring estimates are within the range of our data.

Discussion

Across 369 site-years we found that the full phenological distribution of caterpillars has a strong effect on the relationship between hatching date and breeding success of blue tits, with all three of our EMMH causal predictions receiving support. For *prediction 1*, that mean timing of the caterpillar distribution predicts variation in the optimum hatching date, we find strong empirical support both for probability of nest success and the number fledged, thereby providing the first test of a widely made assumption (e.g. Reed et al., 2013; Vedder et al., 2013). Our finding in support of *prediction 2*, that among site-year variation in the maximum height of the caterpillar abundance distribution predicts variation in the height of the blue tit fitness function for both components of reproductive success, demonstrates the value of moving beyond solely timing based representations and tests of the MMH (see also Weir et al., 2026). For *prediction 3*, that the duration of caterpillar availability should predict the width of the fitness function, we found weaker support, with an effect only apparent for probability of success. As far as we are aware, this is the first empirical demonstration of a theoretical expectation (Johansson et al., 2015; Visser & Gienapp, 2019) that the width of the resource distribution should impact on consumer fitness.

We demonstrate that impacts of caterpillars on the fitness function translate into consequences for blue tit populations, with average hatch dates lagging well behind optimum hatch dates, a lag that is most pronounced when the caterpillar phenology is early. As a result of this lag, for most site-years we estimate directional selection for earlier hatch dates - in broad agreement with studies on selection on avian lay dates (Cao et al., 2019; Charmantier et al., 2008; Radchuk et al., 2019; Visser et al., 2015) – and consequently mean breeding success is depressed well below the predicted local maximum. We also establish that the effects of the height and width of the caterpillar peak on the fitness function translate into impacts on population mean fitness.

Prediction 1 Our analyses demonstrate clearly that timing matters for blue tit breeding success, as has been shown in other studies on tits (Charmantier et al., 2008; Reed et al., 2013) and North American breeding birds (Youngflesh et al., 2023), with substantial decline in fitness for individuals that are 10 or more days either side of the optimum. Most previous observational studies examining the impact of resource-consumer asynchrony on consumer fitness have assumed, either explicitly (e.g. Vedder et al., 2013) or implicitly (e.g. Reed et al., 2013), that resource mean timing determines the consumer optimum (i.e. a slope = 1). Here, we relax and test this assumption, finding a strong positive correlation between the two, with a slope of the optimum hatch date on the mean caterpillar timing that does not depart significantly from 1 for either fitness component. Across the observed range of caterpillar timings, we find that breeding success is maximised when hatch date is 12 (for an early caterpillar site-year) to 23 (for a late caterpillar site-year) days prior to the timing of maximum caterpillar abundance, in broad agreement with earlier work that estimated the relationship between tit consumer-resource asynchrony and fitness (Reed et al., 2013). For the average site-year combination we find that mean hatch dates lag 7-8 days behind the optimum hatch dates and, consistent with other work (e.g., (Charmantier et al., 2008), this lag is exacerbated for site-years with the earliest caterpillar peaks. One explanation for the lag we observe is that climate change has already pushed the system out of equilibrium, though long-term studies on other tit and flycatcher systems show little evidence for the sustained strengthening of directional selection on timing that is expected if lag were generally increasing (Visser et al., 2015, 2021). Another possibility is that the lag arises via trade-offs between phenological optima among fitness components (Cao et al., 2019; Visser & Gienapp, 2019), though Reed et al., (2013) report quite similar impacts of resource-consumer asynchrony on fledging success, recruitment and adult survival. While the average lag among site-years is estimated to be positive, within site-years some nests are predicted to have a negative lag, and for these site-years our model predicts stabilising selection on hatch date.

Prediction 2 Under the EMMH we predicted that the height of the resource distribution should have a causal effect on the maximum of the consumer fitness function. This prediction received strong empirical support, with the height of the caterpillar distribution explaining a substantial proportion of the among site-year variance in the blue tit fitness maxima. Similar positive impacts of resource availability on components of consumer fitness have been shown across a variety of systems (e.g. Durant et al., 2005; Visser et al., 2015; Weir et al., 2026). We find clear evidence for diminishing returns of caterpillar abundance on maximum number fledging (Fig 2c, S2), with the slope of the log-log relationship estimated to be substantially below 1. Here our finding, in broad agreement with estimates obtained for UK populations of blue tit and great tit that adopted a different modelling approach and reported significant positive effects on fledging number and positive but non-significant effects on the probability of success (Weir et al., 2026). With the intention of allowing the impact of resource availability to be dependent on the degree of asynchrony, Weir et al., (2026) include an interaction between asynchrony and resource height, which introduces undesired quadratic relationship between height and maximum fitness (as discussed in Macphie et al., 2023). In comparison, an advantage of our non-linear modelling approach is that we have a term that directly tests our hypothesis for an association between caterpillar height and maximum fitness, whilst allowing the benefit of additional resources to depend on the degree of asynchrony. Through comparison of site-year estimates of the observed hatch dates and fitness functions, we find that a greater maximum availability of caterpillars translates into substantial benefits for mean fitness for asynchronous populations and both fitness components. Our results suggest that attempts to predict impacts of climate-change on consumers would benefit from accounting for changes in resource availability as well as timing (e.g., Macphie et al., 2023).

623 ***Prediction 3*** We provide the first evidence in support of the prediction that the width of the
624 resource phenological distribution has a positive effect on the width of the consumer fitness
625 function (i.e. strength of stabilising selection) (Johansson et al., 2015; Visser & Gienapp, 2019).
626 In the fitness function model, the Poisson response captures more information and therefore
627 should offer more power than the binomial response. Therefore, on solely statistical grounds, it
628 was surprising that the effect of resource width on consumer fitness function width was only
629 significant for the binomial probability of success part of the model. We tentatively suggest that
630 when the caterpillar peak is especially brief, this may tend to elevate levels of brood desertion
631 and complete failure, rather than having negative impacts through brood reduction. In general,
632 we see weaker support for EMMH prediction 3, which is likely to be at least partially attributable
633 to estimates of variances being less precise than estimates of means, in that the width
634 parameter represents the variance of the phenological fitness function. The power to detect an
635 effect of resource distribution width will also be reduced by the strong negative covariance
636 between the height and width of the distribution (Table S1). Both caterpillar height and width
637 were included as predictors of the fitness function width as both parameters influence the
638 duration of time over which the abundance falls above a given threshold, but in comparison to
639 the height and mean timing, the duration of the peak varies rather little.
640
641 Under the classic MMH framing, an increase in asynchrony between bird breeding phenology
642 and the caterpillar mean timing under warmer spring conditions is expected to result in reduced
643 fledging success as birds lag further behind (Visser et al., 2006). If we move to the EMMH
644 framing, in addition to impacting on caterpillar timings, higher spring temperatures are
645 associated with a higher peak of caterpillar abundance (Macphie et al., 2023); but note that
646 other aspects of climatic change may reduce moth populations (Martay et al., 2017). Therefore,
647 holding lag constant, our model predictions suggest that warmer springs may result in higher
648 fledging success, for both the probability of success and the number of offspring fledged. The

649 extent to which this positive impact of warming on caterpillar abundance compensates (or
650 buffers; Weir & Phillimore, 2024) the negative impact that increasing trophic asynchrony is
651 expected to have on fledging success remains to be established. However, our finding that
652 caterpillar abundance has diminishing returns for both fitness components, means that the
653 extent of this buffering effect is likely to be short-lived and may be exhausted if temperatures
654 continue to rise. On a landscape scale, spatial heterogeneity in the caterpillar phenological
655 distribution (Macphie et al., 2025), degree of blue tit-caterpillar asynchrony and impacts of the
656 caterpillar phenological distribution on blue tit fitness, coupled with potential dispersal from
657 matched to mismatched sites via temporally varying source-sink dynamics, may buffer
658 consumer metapopulations (Weir & Phillimore, 2024). Here, we find support for some of the
659 conditions required for such spatial portfolio effects to operate, in that we estimate substantial
660 site-year variance in the maximum abundance of caterpillars, the degree of resource-consumer
661 lag and in the predicted mean fitness.
662
663 The framework for testing the EMMH that we present here has broad applicability to other
664 systems where fitness outcomes are expected to be sensitive to the phenology of interacting
665 species, such as plant-pollinator and host-parasite systems. Our framework also presents clear
666 opportunities for future extensions. For instance, while temperature is widely examined as a
667 driver of phenology (Thackeray et al., 2016), we are aware of only one study that has sought to
668 estimate the effects of temperature on the consumer phenological fitness function (Chevin et al.,
669 2015). Chevin et al., (2015) estimate the thermal sensitivity of optimum timing (*B*) in the Hoge
670 Veluwe population of great tits to be -5.01 days $^{o}C^{-1}$, which combines both direct and indirect
671 (via the resource peak) effects of spring temperatures on the optimum. We can estimate the
672 indirect portion of *B* for our system as the product of (i) the effect of spring temperature on
673 caterpillar timing (-4.96 days $^{o}C^{-1}$, 26) and (ii) the effect of caterpillar timing on the optimum
674 hatch date (0.64 and 0.82 days $^{o}C^{-1}$ for success and number fledged, respectively). This results

675 in estimates of B = -3.17 days $^{o}C^{-1}$ and -4.07 days $^{o}C^{-1}$ for success and number fledged,
676 respectively, suggesting that the thermal sensitivity of optimum timing is largely driven by shifts
677 in caterpillar timing. The impact of spring temperatures on the parameters governing the fitness
678 function are key to predicting population fates in a changing climate (Chevin et al., 2010).
679 Therefore, a priority for future work will be to extend our model to estimate both direct and
680 indirect effects of spring temperature on the consumer fitness function.
681
682 In developing this framework, we have focused on the reproductive success components of
683 fitness, and it is possible that different impacts may be observed on different fitness
684 components. The framework we have developed here could be applied to both recruitment and
685 adult survival, though within our focal system annual recruitment is too infrequent to obtain
686 useful estimates. We opted to focus on fledging success as based on earlier work we expected
687 this fitness component to be especially sensitive to mismatch (Reed et al., 2013). However,
688 fledging success can be seen as the product of investment in clutch size – which will usually be
689 made substantially earlier than the caterpillar peak – and the proportion of those eggs that
690 fledge. The extent to which the relationships we observe are due to each of these processes
691 remains to be established. Interestingly, (Perrins, 1991) found a positive correlation between
692 caterpillar abundance and great tit clutch size across years, suggesting that adult birds may be
693 able to adjust their reproductive investment based on a prediction of future resource availability.
694
695 In the wider evolutionary ecology context, our study is one of few empirical demonstrations that
696 the traits of a resource taxon affect the fitness function for a consumer trait (Schulter & Grant,
697 1984), and the first to demonstrate this in the context of phenology. Meta-analyses have
698 estimated the degree to which selection on traits in general is heterogenous over time
699 (Morrissey & Hadfield, 2012) and/or space (Siepielski et al., 2017). Morrissey & Hadfield (2012)
700 argued that our understanding of the drivers of selection would benefit from work to establish

ecological mechanisms. Here we have identified a clear biological mechanism whereby spatiotemporal variation in spring temperatures has the potential to generate interannual and spatial heterogeneity in both the strength and direction of selection.

In this study, we present an extension to the classic match/mismatch hypothesis, in which the full phenological distribution of a resource is expected to be predictive of the full consumer phenological fitness function. We demonstrate that for blue tits in relation to their caterpillar resource, the resource phenological distribution impacts the consumer phenological fitness function, not only via its mean timing, but also through its height and width. Higher resource abundance increases reproductive success, but the distribution of a resource throughout the reproductive period also impacts on the severity of the negative impacts of asynchrony on fitness. In a warming world, where trophic interactions are prone to uncoupling via unequal responses, accurate estimates of the fitness repercussions of trophic asynchrony are essential for predicting the long-term consequences for populations. Our results suggest that the classic MMH, which focuses primarily on the effects of resource timing, can miss important contributions that changes in the full resource phenological distribution can make to trends in population mean fitness and selection on timing.

Acknowledgments

We are extremely grateful to everyone that has supported data collection over the years, with particular thanks to Jack Shutt, Kat Keogan, Jelmer Samplonius, Ellie Mayhew-Macphie and Megan Stamp for field season management, and to all of the landowners and managers of our field sites. This work was funded by a Natural Environment Research Council (NERC) advanced fellowship (NE/I020598/1O) and NERC grant no. (NE/P011802/1) and a NERC E3 DTP PhD to K.H.M.

References

Bialy, I., Shutt, J., Stamp, M., Phillimore, A., & Macphie, K. (2026). The spring phenology of arboreal beetles and spiders and their potential as alternative prey for birds mistimed with the caterpillar peak. *Authorea*. https://doi.org/10.22541/AUTHOREA.15003146/V1

Both, C., van Asch, M., Bijlsma, R. G., van den Burg, A. B., & Visser, M. E. (2009). Climate change and unequal phenological changes across four trophic levels: constraints or adaptations? *Journal of Animal Ecology*, *78*(1), 73–83. https://doi.org/10.1111/j.1365-2656.2008.01458.x

Burgess, M. D., Smith, K. W., Evans, K. L., Leech, D., Pearce-Higgins, J. W., Branston, C. J., Briggs, K., Clark, J. R., du Feu, C. R., Lewthwaite, K., Nager, R. G., Sheldon, B. C., Smith, J. A., Whytock, R. C., Willis, S. G., & Phillimore, A. B. (2018). Tritrophic phenological match–mismatch in space and time. *Nature Ecology & Evolution*, *2*(6), 970–975. https://doi.org/10.1038/s41559-018-0543-1

Bürkner, P. C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*. https://doi.org/10.18637/JSS.V080.I01

Cao, Y., Visser, M. E., & Tufto, J. (2019). A time-series model for estimating temporal variation in phenotypic selection on laying dates in a Dutch great tit population. *Methods in Ecology and Evolution*, *10*(9), 1401–1411. https://doi.org/10.1111/2041-210X.13249

Charmantier, A., McCleery, R. H., Cole, L. R., Perrins, C., Kruuk, L. E. B., & Sheldon, B. C. (2008). Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, *320*(5877), 800–803. https://doi.org/10.1126/science.1157174

Chevin, L.-M., Lande, R., & Mace, G. M. (2010). Adaptation, Plasticity, and Extinction in a Changing Environment: Towards a Predictive Theory. *PLoS Biology*, *8*(4), e1000357. https://doi.org/10.1371/journal.pbio.1000357

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

Cohen, J. M., Lajeunesse, M. J., & Rohr, J. R. (2018). A global synthesis of animal phenological responses to climate change. *Nature Climate Change*, *8*(3), 224–228. https://doi.org/10.1038/s41558-018-0067-3

Consul, P. C., & Jain, G. C. (1973). A generalization of the poisson distribution. *Technometrics*, *15*(4), 791–799. https://doi.org/10.1080/00401706.1973.10489112

Cushing, D. H. (1969). The Regularity of the Spawning Season of Some Fishes. *ICES Journal of Marine Science*, *33*(1), 81–92. https://doi.org/10.1093/icesjms/33.1.81

Cushing, D. H. (1982). *Climate and fisheries*. Academic Press, London. https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&q=Cushing+DH+1982+Climate+and+fisheries&btnG=

Cushing, D. H. (1990). Plankton production and year-class strength in fish populations: An update of the match/mismatch hypothesis. *Advances in Marine Biology*, *26*(C), 249–293. https://doi.org/10.1016/S0065-2881(08)60202-3

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

Durant, J. M., Hjermann, D. O., Anker-Nilssen, T., Beaugrand, G., Mysterud, A., Pettorelli, N., & Stenseth, N. Chr. (2005). Timing and abundance as key mechanisms affecting trophic interactions in variable environments. *Ecology Letters*, *8*(9), 952–958. https://doi.org/10.1111/j.1461-0248.2005.00798.x

IPCC. (2021). Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]. *Cambridge University Press. In Press.*

Johansson, J., Kristensen, N. P., Nilsson, J. Å., & Jonzén, N. (2015). The eco-evolutionary consequences of interspecifi c phenological asynchrony-a theoretical perspective. *Oikos*, *124*, 102–112. https://doi.org/10.1111/oik.01909

Kharouba, H. M., Ehrlén, J., Gelman, A., Bolmgren, K., Allen, J. M., Travers, S. E., & Wolkovich, E. M. (2018). Global shifts in the phenological synchrony of species interactions over recent decades. *Proceedings of the National Academy of Sciences*, *115*(20), 5211–5216. https://doi.org/10.1073/PNAS.1714511115

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., & Arnold, S. J. (1983). The measurement of selection on correlated characters. *Evolution*, *37*(6), 1210–1226.

Macphie, K. H. (2023). *The full phenological distribution and the match/mismatch hypothesis*.

Macphie, K. H., & Phillimore, A. B. (2024). Phenology. *Current Biology*, *34*(5), R183–R188. https://doi.org/10.1016/j.cub.2024.01.007

Macphie, K. H., Samplonius, J. M., Hadfield, J. D., Pearce Higgins, J. W., & Phillimore, A. B. (2025). Tree taxon effects on the phenology of caterpillar abundance and biomass. *Oikos*, *2025*(4), e10972. https://doi.org/10.1111/OIK.10972

Macphie, K. H., Samplonius, J. M., Pick, J. L., Hadfield, J. D., & Phillimore, A. B. (2023). Modelling thermal sensitivity in the full phenological distribution: A new approach applied to the spring arboreal caterpillar peak. *Functional Ecology*, *37*(12), 3015–3026. https://doi.org/10.1111/1365-2435.14436

Martay, B., Brewer, M. J., Elston, D. A., Bell, J. R., Harrington, R., Brereton, T. M., Barlow, K. E., Botham, M. S., & Pearce-Higgins, J. W. (2017). Impacts of climate change on national biodiversity population trends. *Ecography*, *40*(10), 1139–1151. https://doi.org/10.1111/ECOG.02411

Morrissey, M. B., & Hadfield, J. D. (2012). DIRECTIONAL SELECTION IN TEMPORALLY REPLICATED STUDIES IS REMARKABLY CONSISTENT. *Evolution*, *66*(2), 435–442. https://doi.org/10.1111/J.1558-5646.2011.01444.X

Parmesan, C., & Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, *421*, 37–42. www.nature.com/nature

PERRINS, C. M. (1991). Tits and their caterpillar food supply. *Ibis*, *133*, 49–54. https://doi.org/10.1111/J.1474-919X.1991.TB07668.X

Phillimore, A. B., Leech, D. I., Pearce-Higgins, J. W., & Hadfield, J. D. (2016). Passerines may be sufficiently plastic to track temperature-mediated shifts in optimum lay date. *Global Change Biology*, *22*(10), 3259–3272. https://doi.org/10.1111/GCB.13302

Probst, C. M., Yanco, S., Clark, I., Ziebell, M., Fuirst, M., Mackenzie, S. A., Ibáñez, I., & Weeks, B. C. (2026). Resource declines shape phenological and morphological responses to climate change. *Proceedings of the National Academy of Sciences*, *123*(26), e2607714123. https://doi.org/10.1073/PNAS.2607714123

Radchuk, V., Reed, T., Teplitsky, C., van de Pol, M., Charmantier, A., Hassall, C., Adamík, P., Adriaensen, F., Ahola, M. P., Arcese, P., Miguel Avilés, J., Balbontin, J., Berg, K. S., Borras, A., Burthe, S., Clobert, J., Dehnhard, N., de Lope, F., Dhondt, A. A., … Kramer-Schadt, S. (2019). Adaptive responses of animals to climate change are most likely insufficient. *Nature Communications*, *10*(1), 1–14. https://doi.org/10.1038/s41467-019-10924-4

Ramakers, J. J. C., Gienapp, P., & Visser, M. E. (2019). Comparing two measures of phenological synchrony in a predator–prey interaction: Simpler works better. *Journal of Animal Ecology*, *00*, 1–12. https://doi.org/10.1111/1365-2656.13143

Reed, T. E., Grøtan, V., Jenouvrier, S., Sæther, B. E., & Visser, M. E. (2013). Population Growth in a Wild Bird Is Buffered Against Phenological Mismatch. *Science*, *340*(6131), 488–491. https://doi.org/10.1126/science.1232870

Roslin, T., Antão, L., Hällfors, M., Meyke, E., Lo, C., Tikhonov, G., Delgado, M. del M., Gurarie, E., Abadonova, M., Abduraimov, O., Adrianova, O., Akimova, T., Akkiev, M., Ananin, A., Andreeva, E., Andriychuk, N., Antipin, M., Arzamascev, K., Babina, S., … Ovaskainen, O. (2021). Phenological shifts of abiotic events, producers and consumers across a continent. *Nature Climate Change*, *11*(3), 241–248. https://doi.org/10.1038/s41558-020-00967-7

Samplonius, J. M., Atkinson, A., Hassall, C., Keogan, K., Thackeray, S. J., Assmann, J. J., Burgess, M. D., Johansson, J., Macphie, K. H., Pearce-Higgins, J. W., Simmonds, E. G., Varpe, Ø., Weir, J. C., Childs, D. Z., Cole, E. F., Daunt, F., Hart, T., Lewis, O. T., Pettorelli, N., … Phillimore, A. B. (2020). Strengthening the evidence base for temperature-mediated phenological asynchrony and its impacts. *Nature Ecology & Evolution*, *5*(2), 155–164. https://doi.org/10.1038/s41559-020-01357-0

Schulter, D., & Grant, P. R. (1984). Determinants of Morphological Patterns in Communities of Darwin's Finches on JSTOR. *The American Naturalist*, *123*(2), 175–196. https://www.jstor.org/stable/2461032?seq=20

Shutt, J. D., Bolton, M., Cabello, I. B., Burgess, M. D., & Phillimore, A. B. (2018). The effects of woodland habitat and biogeography on blue tit (Cyanistes caeruleus) territory occupancy and productivity along a 220km transect. *Ecography*, 1–12. https://doi.org/10.1111/ecog.03573

Shutt, J. D., Burgess, M. D., & Phillimore, A. B. (2019). A spatial perspective on the phenological distribution of the spring woodland caterpillar peak. *The American Naturalist*, *194*, E109–E121.

Siepielski, A. M., Morrissey, M. B., Buoro, M., Carlson, S. M., Caruso, C. M., Clegg, S. M., Coulson, T., DiBattista, J., Gotanda, K. M., Francis, C. D., Hereford, J., Kingsolver, J. G., Augustine, K. E., Kruuk, L. E. B., Martin, R. A., Sheldon, B. C., Sletvold, N., Svensson, E. I., Wade, M. J., & MacColl, A. D. C. (2017). Precipitation drives global variation in natural selection. *Science*, *355*(6328), 959–962. https://doi.org/10.1126/SCIENCE.AAG2773/SUPPL_FILE/SIEPIELSKI.SM_CORRECTED.PDF

Stan Development Team. (2022). *RStan: the R interface to Stan. R package version 2.26.13. http://mc-stan.org/*.

Thackeray, S. J., Henrys, P. A., Hemming, D., Bell, J. R., Botham, M. S., Burthe, S., Helaouet, P., Johns, D. G., Jones, I. D., Leech, D. I., Mackay, E. B., Massimino, D., Atkinson, S., Bacon, P. J., Brereton, T. M., Carvalho, L., Clutton-Brock, T. H., Duck, C., Edwards, M., … Wanless, S. (2016). Phenological sensitivity to climate across taxa and trophic levels. *Nature*, *535*(7611), 241–245. https://doi.org/10.1038/nature18608

Thackeray, S. J., Sparks, T. H., Frederiksen, M., Burthe, S., Bacon, P. J., Bell, J. R., Botham, M. S., Brereton, T. M., Bright, P. W., Carvalho, L., Clutton-Brock, T., Dawson, A., Edwards, M., Elloitt, J. M., Harrington, R., Johns, D., Jones, I. D., Jones, J. T., Leech, D. I., … Wanless, S. (2010). Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology*, *16*(12), 3304–3313. https://doi.org/10.1111/j.1365-2486.2010.02165.x

Vedder, O., Bouwhuis, S., & Sheldon, B. C. (2013). Quantitative Assessment of the Importance of Phenotypic Plasticity in Adaptation to Climate Change in Wild Bird Populations. *PLoS Biology*, *11*(7), e1001605. https://doi.org/10.1371/journal.pbio.1001605

Visser, M. E., & Gienapp, P. (2019). Evolutionary and demographic consequences of phenological mismatches. *Nature Ecology & Evolution*, *3*, 879–885. https://doi.org/10.1038/s41559-019-0880-8

Visser, M. E., Gienapp, P., Husby, A., Morrisey, M., de la Hera, I., Pulido, F., & Both, C. (2015). Effects of Spring Temperatures on the Strength of Selection on Timing of Reproduction in a Long-Distance Migratory Bird. *PLOS Biology*, *13*(4), e1002120. https://doi.org/10.1371/JOURNAL.PBIO.1002120

Visser, M. E., Holleman, L. J. M., & Gienapp, P. (2006). Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia*, *147*(1), 164–172. https://doi.org/10.1007/s00442-005-0299-6

Visser, M. E., Lindner, M., Gienapp, P., Long, M. C., & Jenouvrier, S. (2021). Recent natural variability in global warming weakened phenological mismatch and selection on seasonal timing in great tits (Parus major). *Proceedings of the Royal Society B*, *288*(1963). https://doi.org/10.1098/RSPB.2021.1337

Weir, J. C., & Phillimore, A. B. (2024). Buffering and phenological mismatch: A change of perspective. *Global Change Biology*, *30*(5), e17294. https://doi.org/10.1111/GCB.17294

Weir, J. C., Smith, K. W., Smith, L., Briggs, K., Clark, J. R., Feu, C. R. du, Flack, I., Thomson, L., Wilton, B., & Burgess, M. D. (2026). *Resource abundance can buffer trophic mismatch in a caterpillar-passerine food-chain*. https://doi.org/10.32942/X2KD4J

Youngflesh, C., Montgomery, G. A., Saracco, J. F., Miller, D. A. W., Guralnick, R. P., Hurlbert, A. H., Siegel, R. B., LaFrance, R., & Tingley, M. W. (2023). Demographic consequences of phenological asynchrony for North American songbirds. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(28), e2221961120. https://doi.org/10.1073/PNAS.2221961120/SUPPL_FILE/PNAS.2221961120.SAPP.PDF

Supporting Information: Timing isn't everything: impacts of maximum abundance and duration of a seasonal resource on consumer fitness

Table S1: Fitness model summary.

Figure S1: Effect of caterpillar distribution height on blue tit fitness function width.

Figure S2: Cropped x-axis showing the effect of caterpillar distribution height on the fitness function maximum.

Figure S3: Posterior distributions for the proportion of the variance in population mean fitness and directional selection explained by the caterpillar distribution.

Figure S4: Mean predicted trend in probability of success and the number fledged by hatch date.

Model fit to data

Figure S5: Histograms of the total number of young fledged, the variance in number fledged and proportion of nests that failed; predicted from simulations under the model across the posterior distribution.

Table S1: Fitness model summary. Coefficient estimates, posterior distribution summary, effective sample sizes, Rhat and the prior distribution standard deviation or scale (see main text). Correlation matrices used to estimate covariances were consistent across site, year and site-year terms; there is no prior distribution information because the prior was put on the Cholesky factor. Effective sample sizes and Rhat were used to ensure sufficient sampling and model convergence; they are not available for some components of the correlation matrices as the value is fixed at one. SD = standard deviation.

Coefficient	Posterior Mean	Posterior Median	95% Credible Interval		Effective Sample Size	Rhat	Prior Standard Deviation or Scale
			2.5%	97.5%			
Probability of Success - Bernoulli							
θ_S intercept	-8.812	-8.16	-17.59	-3.863	3197	1.000	20
$ln\omega_S$ intercept	2.525	2.481	2.061	3.208	2091	1.000	5
$f_S M_S$ intercept	4.331	4.268	2.841	6.177	2498	1.001	$\sqrt{\pi^2/3}$
θ_S: θ_C slope	0.643	0.640	0.250	1.059	3144	1.001	5
$ln\omega_S$: $ln\omega_C$ slope	1.530	1.446	-0.089	3.745	1641	1.003	5
$ln\omega_S$: $f_C M_C$ slope	0.210	0.189	-0.074	0.619	1945	1.001	5
$f_S M_S$: $f_C M_C$ slope	0.698	0.678	0.233	1.271	2478	1.002	5
θ_S: Site SD	1.359	0.984	0.042	4.831	3972	1.000	10
$ln\omega_S$: Site SD	0.190	0.154	0.008	0.570	3392	1.000	3
$f_S M_S$: Site SD	0.689	0.66	0.078	1.485	2780	1.000	3
θ_S: Year SD	1.698	1.232	0.057	5.850	3825	1.000	10
$ln\omega_S$: Year SD	0.166	0.124	0.006	0.571	3366	1.000	3
$f_S M_S$: Year SD	0.910	0.846	0.272	1.912	3562	1.000	3
θ_S: Site-year SD	1.550	1.180	0.051	5.001	2564	1.003	10
$ln\omega_S$: Site-year SD	0.279	0.258	0.013	0.697	2473	1.000	3
$f_S M_S$: Site-year SD	0.455	0.388	0.019	1.242	3155	1.001	3
$f_S M_S$: Female SD	2.434	2.408	0.904	4.062	1982	1.001	10
Correlation matrix: θ_S: θ_S	1.000	1.000	1.000	1.000	*NA*	*NA*	*NA*
$ln\omega_S$: θ_S	-0.184	-0.220	-0.854	0.653	3424	1.000	*NA*
$f_S M_S$: θ_S	-0.010	-0.008	-0.756	0.749	2778	1.000	*NA*
θ_S: $ln\omega_S$	-0.184	-0.220	-0.854	0.653	3424	1.000	*NA*
$ln\omega_S$: $ln\omega_S$	1.000	1.000	1.000	1.000	1015	0.999	*NA*
$f_S M_S$: $ln\omega_S$	-0.108	-0.141	-0.789	0.688	3173	1.001	*NA*
θ_S: $f_S M_S$	-0.010	-0.008	-0.756	0.749	2778	1.000	*NA*
$ln\omega_S$: $f_S M_S$	-0.108	-0.141	-0.789	0.688	3173	1.001	*NA*
$f_S M_S$: $f_S M_S$	1.000	1.000	1.000	1.000	3923	0.999	*NA*

Coefficient	Posterior Mean	Posterior Median	95% Credible Interval		Effective Sample Size	Rhat	Prior Standard Deviation or Scale
			2.5%	97.5%			
Number fledged (conditional on success) – Truncated Generalised Poisson							
θ_N intercept	-9.757	-9.356	-16.26	-5.274	2967	1.000	20
$ln\omega_N$ intercept	3.124	3.111	2.857	3.471	3111	1.000	5
$f_N M_N$ intercept	1.952	1.952	1.856	2.042	3654	1.001	10
θ_N: θ_C slope	0.812	0.817	0.485	1.121	3467	1.000	5
$ln\omega_N$: $ln\omega_C$ slope	-0.018	-0.014	-1.09	1.027	1550	1.004	5
$ln\omega_N$: $f_C M_C$ slope	-0.090	-0.087	-0.294	0.104	1962	1.003	5
$f_N M_N$: $f_C M_C$ slope	0.127	0.127	0.087	0.167	3213	1.000	5
θ_N: Site SD	1.875	1.582	0.084	5.418	3789	1.000	10
$ln\omega_N$: Site SD	0.212	0.201	0.013	0.493	3397	1.000	3
$f_N M_N$: Site SD	0.094	0.094	0.038	0.149	3563	1.000	3
θ_N: Year SD	3.036	2.782	0.158	7.838	3513	1.000	10
$ln\omega_N$: Year SD	0.148	0.119	0.006	0.453	3783	1.000	3
$f_N M_N$: Year SD	0.070	0.067	0.006	0.156	3254	1.002	3
θ_N: Site-year SD	3.347	3.371	0.286	6.843	1894	1.000	10
$ln\omega_N$: Site-year SD	0.197	0.187	0.013	0.456	1959	1.000	3
$f_N M_N$: Site-year SD	0.042	0.039	0.003	0.097	1979	1.001	3
$f_N M_N$: Female SD	0.126	0.128	0.064	0.168	2886	1.002	10
Over/under dispersion	-0.047	-0.047	-0.056	-0.038	3394	1.000	10
Correlation matrix: θ_N: θ_N	1.000	1.000	1.000	1.000	*NA*	*NA*	*NA*
$ln\omega_N$: θ_N	-0.363	-0.438	-0.895	0.516	3035	1.000	*NA*
$f_N M_N$: θ_N	-0.234	-0.274	-0.833	0.571	2734	1.001	*NA*
θ_N: $ln\omega_N$	-0.363	-0.438	-0.895	0.516	3035	1.000	*NA*
$ln\omega_N$: $ln\omega_N$	1.000	1.000	1.000	1.000	197	0.999	*NA*
$f_N M_N$: $ln\omega_N$	-0.102	-0.127	-0.763	0.665	2934	1.000	*NA*
θ_N: $f_N M_N$	-0.234	-0.274	-0.833	0.571	2734	1.001	*NA*
$ln\omega_N$: $f_N M_N$	-0.102	-0.127	-0.763	0.665	2934	1.000	*NA*
$f_N M_N$: $f_N M_N$	1.000	1.000	1.000	1.000	4116	0.999	*NA*
Bernoulli-TGP: θ regression	0.135	0.087	-3.098	3.668	2397	1.000	10
fM regression	0.067	0.060	-0.017	0.192	2451	1.002	10

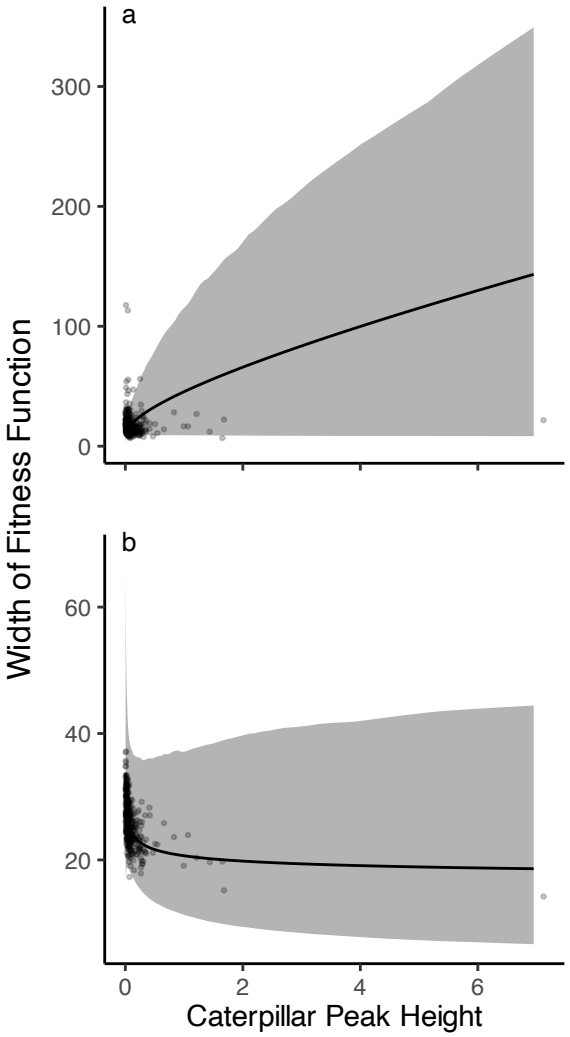
Table continues on next page

920

Coefficient	Posterior Mean	Posterior Median	95% Credible Interval		Effective Sample Size	Rhat	Prior Standard Deviation or Scale
			2.5%	97.5%			
Caterpillar Phenological Distribution - Poisson							
θ_C intercept	8.315	8.294	4.574	12.249	4012	1.000	20
$ln\omega_C$ intercept	2.530	2.527	2.366	2.711	2575	1.001	5
$f_C M_C$ intercept	-2.929	-2.929	-3.573	-2.277	4480	0.999	10
θ_C: Site SD	5.724	5.653	4.337	7.524	2732	1.000	10
$ln\omega_C$: Site SD	0.201	0.199	0.134	0.281	1054	1.003	10
$f_C M_C$: Site SD	0.852	0.842	0.672	1.082	3242	1.000	10
θ_C: Year SD	5.443	5.148	3.300	9.072	3083	1.001	10
$ln\omega_C$: Year SD	0.233	0.221	0.118	0.415	1300	1.003	10
$f_C M_C$: Year SD	0.882	0.837	0.544	1.494	2136	1.001	10
θ_C: Site-year SD	2.516	2.501	1.818	3.283	1165	1.005	10
$ln\omega_C$: Site-year SD	0.152	0.151	0.089	0.222	262	1.009	10
$f_C M_C$: Site-year SD	0.684	0.683	0.607	0.767	2202	1.002	10
$f_C M_C$: Site-year-day SD	0.450	0.450	0.389	0.510	935	1.000	10
$f_C M_C$: TreeID SD	0.595	0.594	0.54	0.651	2290	1.004	10
$f_C M_C$: Recorder SD	0.445	0.438	0.325	0.607	2815	1.002	10
$f_C M_C$: Residual SD	0.828	0.828	0.776	0.879	475	1.007	10
Correlation matrix: $\theta_C : \theta_C$	1.000	1.000	1.000	1.000	*NA*	*NA*	*NA*
$ln\omega_C : \theta_C$	0.174	0.179	-0.146	0.477	485	1.008	*NA*
$f_C M_C : \theta_C$	-0.247	-0.249	-0.442	-0.039	1717	1.001	*NA*
$\theta_C : ln\omega_C$	0.174	0.179	-0.146	0.477	485	1.008	*NA*
$ln\omega_C : ln\omega_C$	1.000	1.000	1.000	1.000	235	0.999	*NA*
$f_C M_C : ln\omega_C$	-0.749	-0.753	-0.873	-0.592	326	1.011	*NA*
$\theta_C : f_C M_C$	-0.247	-0.249	-0.442	-0.039	1717	1.001	*NA*
$ln\omega_C : f_C M_C$	-0.749	-0.753	-0.873	-0.592	326	1.011	*NA*
$f_C M_C : f_C M_C$	1.000	1.000	1.000	1.000	3988	0.999	*NA*

921

922



923

924 Figure S1. Effect of caterpillar distribution height on blue tit fitness function width for the a)

925 probability of success and b) number fledged. Gray shaded areas correspond to the 95%

926 credible intervals.

927

928

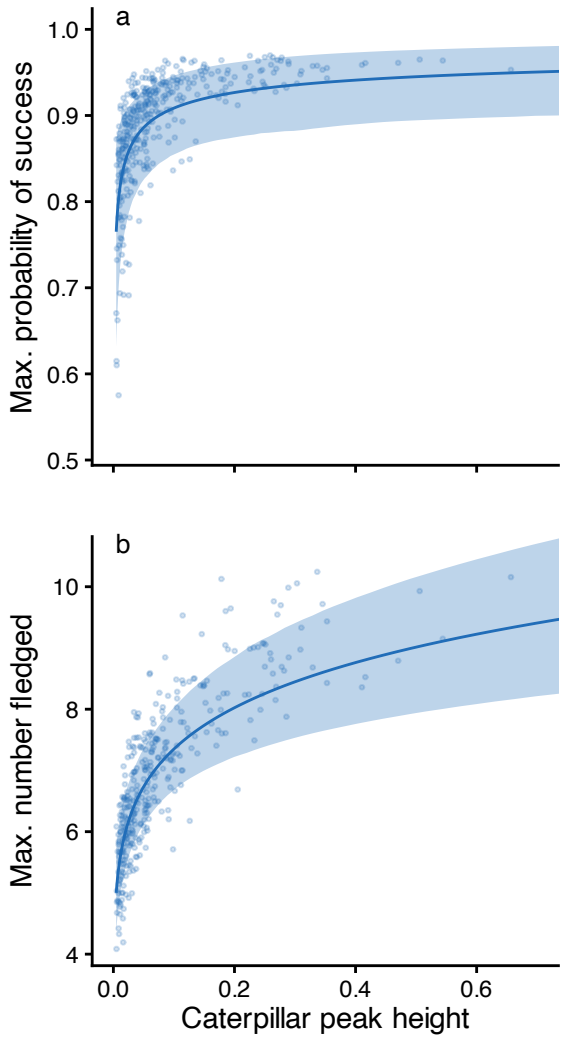
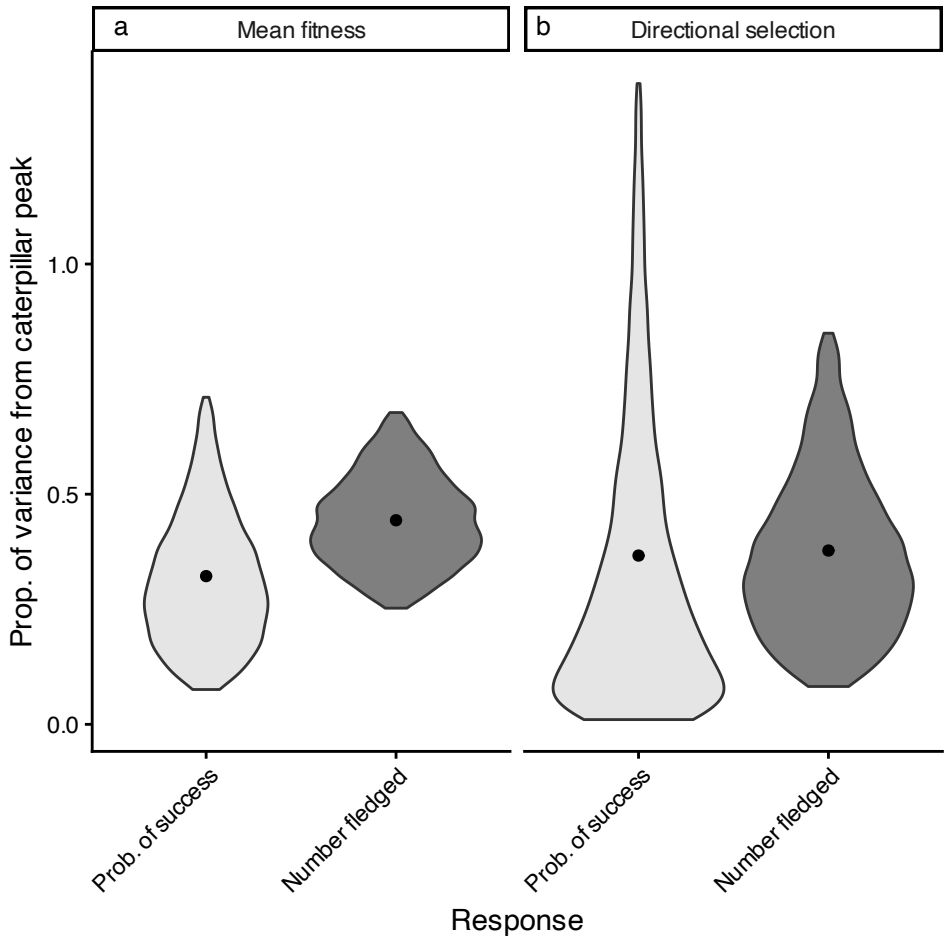


Figure S2. Cropped x-axis showing the effect of caterpillar distribution height on the fitness function maximum for the a) probability of success and b) number fledged. Blue shaded areas correspond to the 95% credible intervals.

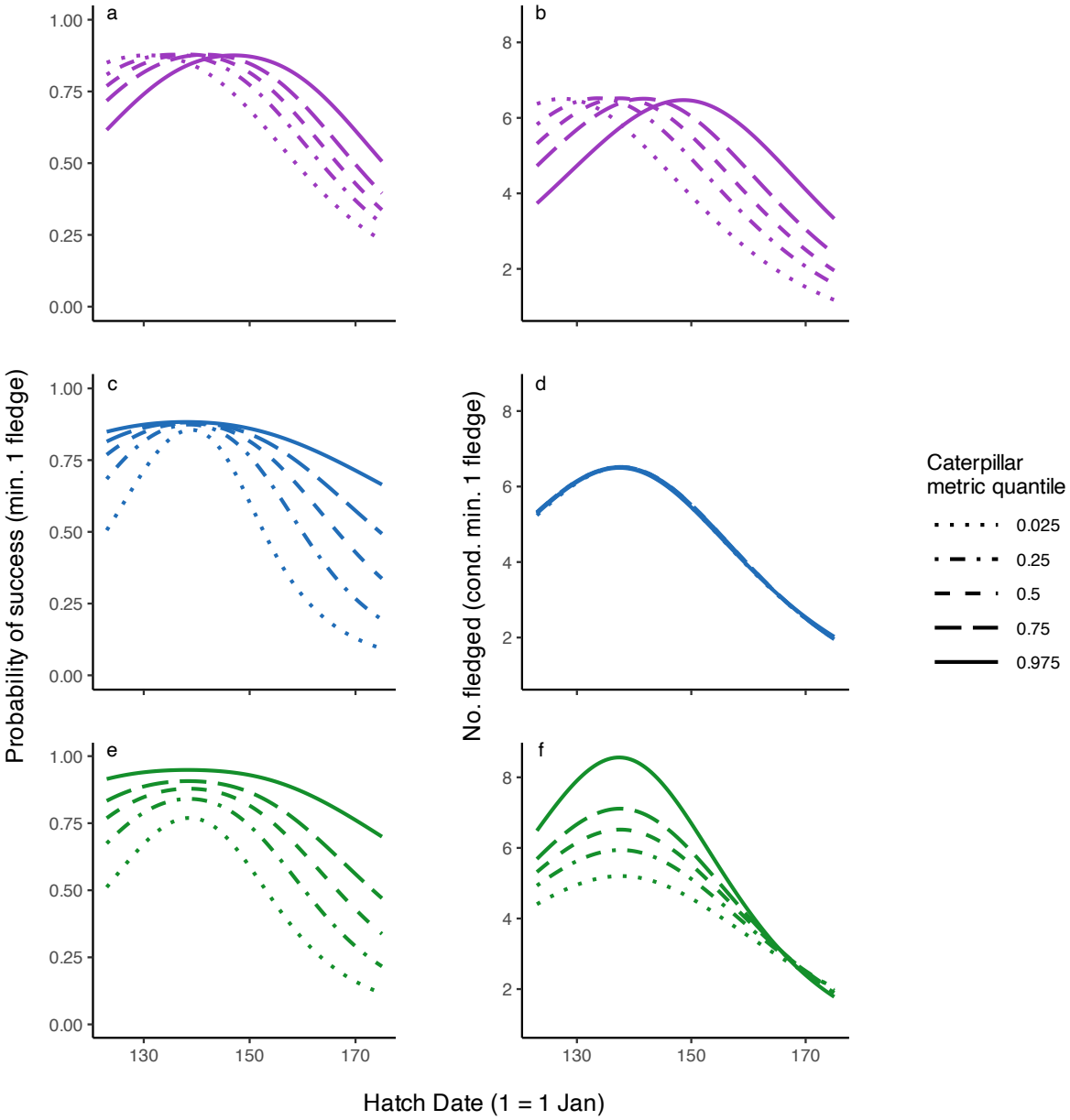


935

936 Figure S3: Posterior distributions for the proportion of the variance in population mean fitness
937 and directional selection explained by the caterpillar distribution. Violins show the 95% credible
938 intervals with the mean indicated by a point.

939

940



941

942 Figure S4. Mean predicted trend in probability of success (a,c,e) and the number fledged (b,d,f)

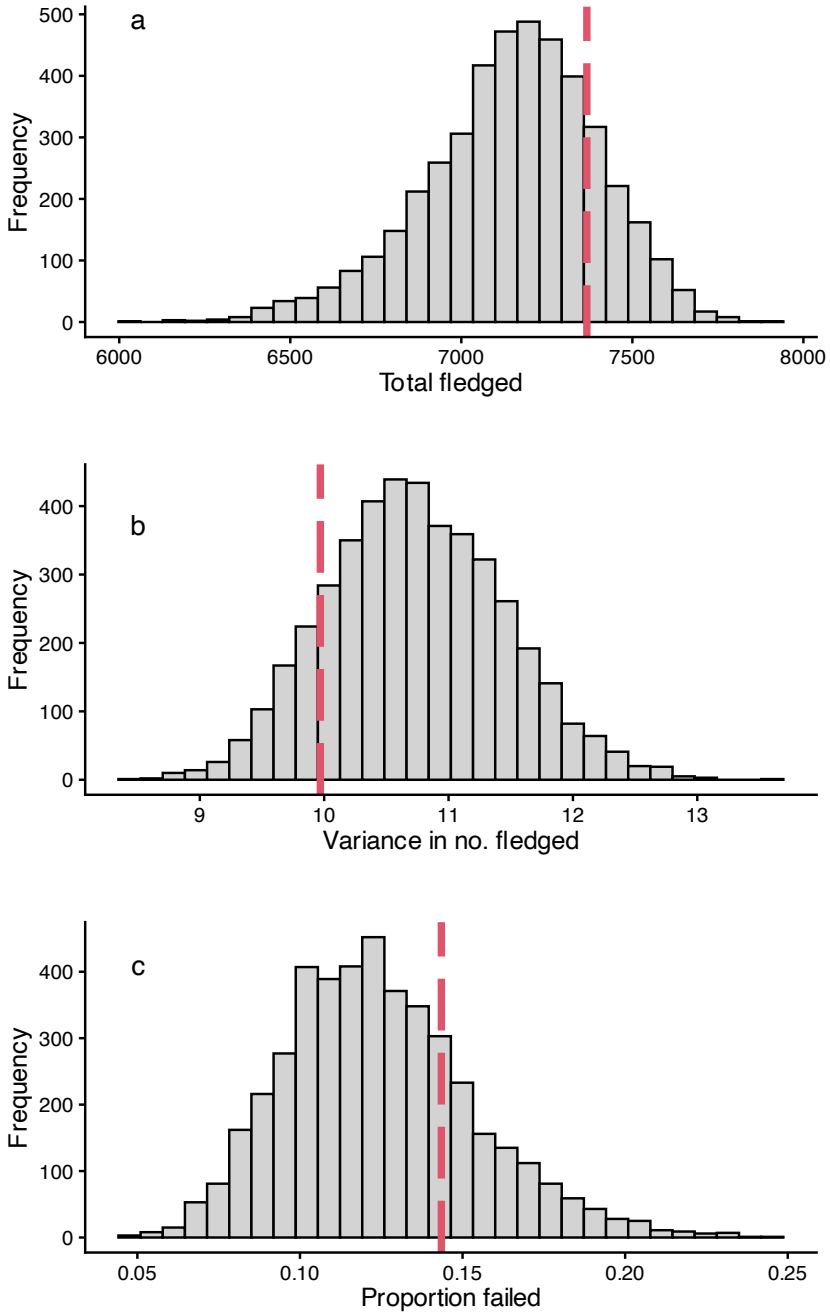
943 by hatch date. Each plot varies one caterpillar metric whilst the others were held at their 0.5

944 quantile: caterpillar mean timing in purple (a,b), height in blue (c,d) and width in green (e,f).

945

Model fit to the data

We used posterior predictive simulates to check model fit. As 10 site-year combinations had not had any successful nests, they were not included in the TGP part of the *Fitness model*. We ran an alternate model with the same composition which also estimated TGP coefficients for the site-year combinations that had no data, allowing simulation across the entire dataset. Point estimates between the *Fitness model* and this alternate model were consistent, but the alternate model had higher uncertainty in some parameter estimates. We compared the distributions of predicted total number of offspring fledged, the variance in number fledged and predicted proportion of nest failures, across all nests in our dataset, to the true values from our data; the true value for each fell well within the estimates across the posterior (Figure S5).



957

958 Figure S5: Histograms of the a) total number of young fledged, b) the variance in number

959 fledged and c) proportion of nests that failed, as predicted from simulations under the model

960 across the posterior distribution. The red dashed lines indicate the true value from our data.