Edinburgh Research Explorer

Passerines may be sufficiently plastic to track temperature-mediated shifts in optimum lay date

Passerines may be sufficiently plastic to track temperature-mediated shifts in optimum lay date

ALBERT B. PHILLIMORE[1], DAVID I. LEECH[2], JAMES W. PEARCE-HIGGINS[2] and JARROD D. HADFIELD[1]

[1]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, UK, [2]British Trust for Ornithology, Thetford Norfolk IP24 2PU, UK

Abstract

Projecting the fates of populations under climate change is one of global change biology's foremost challenges. Here, we seek to identify the contributions that temperature-mediated local adaptation and plasticity make to spatial variation in nesting phenology, a phenotypic trait showing strong responses to warming. We apply a mixed modeling framework to a Britain-wide spatiotemporal dataset comprising >100 000 records of first egg dates from four single-brooded passerine bird species. The average temperature during a specific time period (sliding window) strongly predicts spatiotemporal variation in lay date. All four species exhibit phenological plasticity, advancing lay date by 2–5 days °C^{-1}. The initiation of this sliding window is delayed further north, which may be a response to a photoperiod threshold. Using clinal trends in phenology and temperature, we are able to estimate the temperature sensitivity of selection on lay date (B), but our estimates are highly sensitive to the temporal position of the sliding window. If the sliding window is of fixed duration with a start date determined by photoperiod, we find B is tracked by phenotypic plasticity. If, instead, we allow the start and duration of the sliding window to change with latitude, we find plasticity does not track B, although in this case, at odds with theoretical expectations, our estimates of B differ across latitude vs. longitude. We argue that a model combining photoperiod and mean temperature is most consistent with current understanding of phenological cues in passerines, the results from which suggest that each species could respond to projected increases in spring temperatures through plasticity alone. However, our estimates of B require further validation.

Keywords: citizen science, climate, cues, environmental sensitivity of selection, local adaptation, optimum, phenology, plasticity, space for time

Received 3 September 2015; revised version received 25 March 2016 and accepted 27 March 2016

Introduction

Predicting how populations will respond under different climate change scenarios is a major challenge for ecologists and evolutionists. To take into account both evolutionary and ecological processes, studies typically focus on traits that covary with both climate and fitness (Gienapp *et al.*, 2008), such as breeding phenology in birds. In temperate bird species, lay date often correlates negatively with spring temperatures (Crick & Sparks, 1999) and the timing of hatching impacts upon the fitness of parents and offspring (Visser *et al.*, 2004; Verhulst & Nilsson, 2008).

Chevin *et al.* (2010) presented a model to predict the maximum rate of environmental change a population can cope with. The parameters of the model are maximum population growth rate, generation time, the strength of stabilizing selection, narrow sense heritability, a linear relationship between the environmental variable and the optimum, termed the environmental sensitivity of selection (B), and linear plasticity with respect to the environment (b). A key component of Chevin *et al.*'s equation is $|B - b|$, the degree to which plasticity tracks the phenotypic optimum across environments. If the absolute difference is small, a population is projected to be able to withstand rapid environmental change via plasticity alone, whereas if the difference is large, the persistence of a population in the face of environmental change may depend on evolutionary adaptation. Numerous studies have estimated plasticity of lay date in birds, but very few have quantified B, which depends on quantifying the fitness surface in different environments (Chevin *et al.*, 2015). Recently two studies on wild populations of great tits at single sites have estimated B and parameterized Chevin *et al.*'s model (Gienapp *et al.*, 2013; Vedder *et al.*, 2013). Both studies project that under the Chevin *et al.* model the phenological plasticity of lay date will allow populations to track the optimum phenology even under rapid rates of spring warming.

Correspondence: Albert B. Phillimore, tel. +44 131 650 5413, fax +44 131 650 5455, e-mail: albert.phillimore@ed.ac.uk

For ecological and evolutionary processes that play out over timescales that are not conducive to longitudinal study, space is often substituted for time (Pickett, 1989). Biotic responses to geographic variation in temperature may have arisen over many generations, and therefore be informative about long-term responses to temperature change (Dunne *et al.*, 2004; Phillimore *et al.*, 2010). Consider two populations, where one has experienced average temperatures 1 °C warmer than the other, giving rise to a phenotypic difference of x (as a result of plasticity and/or local adaptation). We can use this difference to predict that a future 1 °C rise in temperature will require the phenotype to change by x (via plasticity and/or microevolution) for the populations to remain as locally adapted as they currently are (Phillimore *et al.*, 2010).

Reciprocal transplants remain the primary tool for detecting local adaptation (Kawecki & Ebert, 2004), but have rarely been attempted on birds due to their mobility. As a result, we know very little about the degree to which geographic variation in lay date represents local adaptation (but see Blondel *et al.*, 1999, 1990). Phillimore *et al.* (2010) presented a statistical approach to estimate local adaptation to an environmental driver that may be useful where reciprocal transplants are unfeasible. This approach relies on the statistical decomposition of spatiotemporal covariation between phenology and temperature data into the contributions of temperature-mediated plasticity and local adaptation (Phillimore *et al.*, 2010). Assuming all populations of a species have the same plastic reaction norm b with respect to a thermal cue, the difference between the slopes of phenology on temperature estimated among locations (spatial slope) vs. among years (temporal slope, assumed to equal b) estimates the contribution made by local adaptation. This method requires that microevolution over a short time period (years to decades) has had little impact on the temporal slope. We suggest that this assumption is reasonable in birds as most empirical studies find substantial plasticity and scant evidence for microevolution (Charmantier & Gienapp, 2013). By contrast, there may have been many generations over which selection could generate locally adapted phenotypes.

When the environment has a deterministic linear trend, such as with latitude, populations are expected to be able to track the optimal reaction norm slope (B) by local adaptation alone (Slatkin, 1978) when population density is constant in space. With stochasticity, the environment may deviate locally from the linear trend and then dispersal will disrupt the ability of populations to attain B. The magnitude of the discrepancy increases as the scale of dispersal becomes long relative to the scale of spatial autocorrelation in the detrended

environment (Hadfield, in press). However, at migration/selection balance, the ratio of the slopes of phenology and temperature each regressed on a linear spatial vector can be used as an estimate of B (Hadfield, in press). Although this method requires more assumptions than direct estimation of B from data on the fitness of individuals (e.g., Chevin *et al.*, 2015), it can be applied to less tractable study systems.

To accurately project population phenological responses into the future requires that we have correctly identified the cue or cues. Unfortunately, the precise cues that birds use to schedule nesting phenology remain uncertain (Caro *et al.*, 2013). Laboratory experiments reveal that increasing day length cues gonadal development (Dawson *et al.*, 2001) and can bring forward the first egg date (Lambrechts *et al.*, 1996), suggesting that photoperiod may set a hard limit before which lay date is unresponsive to other environmental conditions. However, photoperiod cannot account for the year-to-year *in situ* variation in lay date that populations exhibit. Significant regression slopes for lay date on average temperatures during sliding windows have often been used to infer a role for temperature in explaining among year variation (Crick & Sparks, 1999; Husby *et al.*, 2010; Ockendon *et al.*, 2013), with model fit used to identify the time period corresponding to greatest thermal sensitivity. While a negative correlation between temperature and phenology has been observed for many species, there remains a question as to whether temperature has a direct effect as a cue or constraint or an indirect effect, for example, mediated via an impact on food availability (Perrins, 1965). Experiments provide support for both direct and indirect effects, as advances in great tit lay dates can be induced by supplementary feeding in the wild (Robb *et al.*, 2008) and experimental warming in aviaries (Visser *et al.*, 2009), although in each case the induced advance is less than the amount by which a population mean lay date can differ between warm and cold years. Aviary temperature experiments also provide some evidence that great tits may be sensitive to the rate of temperature increase rather than mean temperature (Schaper *et al.*, 2012).

In a spatial context, an additional challenge is that both the cue(s) and plastic responses may vary among populations of a single species, such that the results from one study of a single species may not be transferable to other populations. For instance, studies across populations of blue tits (Porlier *et al.*, 2012) and great tits (Slagsvold, 1976; Husby *et al.*, 2010) have found that the sliding window over which nesting phenology correlates most strongly with temperature differs among populations. For great tits, there is some suggestion that the period of thermal sensitivity begins later and is

shorter in duration as latitude increases from central to northern Europe (Slagsvold, 1976). Extensive spatiotemporal nest record datasets collected by national ornithological societies (e.g., Crick *et al.*, 2003) offer the potential for an examination of spatial variation in the cues responsible for nesting phenology, but have been surprisingly underutilized (but see Gienapp *et al.*, 2010). Replication of phenological observations across years permits analysis of the period of thermal sensitivity, while replication of observations across space permits analysis of whether the sensitive period varies geographically, potentially as a function of photoperiod (Phillimore *et al.*, 2013).

Here, we consider an exceptional spatiotemporal dataset (Crick *et al.*, 2003), comprising in excess of 100 000 observations of first egg dates across four predominantly single-brooded species. We extend our earlier framework for separating the contributions of plasticity and local adaptation (Phillimore *et al.*, 2010, 2012, 2013), to additionally estimate the environmental sensitivity of selection (equivalent to B in Chevin *et al.*, 2010) and the contribution of sustained temporal microevolution during the study period. The aims of the study are threefold: first, to identify the environmental variables that best explain spatiotemporal variation in nesting phenology. Second, to estimate phenological plasticity (b) with respect to thermal cues. Third, to quantify the environmental sensitivity of selection (B) and the ability of plasticity to track the optimum $|B - b|$.

Materials and methods

Data

We focus on four predominantly single-brooded passerine species, defined as those that usually fledge a maximum of one brood per pair per year, for which there are abundant nest record data. Three of these, blue tit (*Cyanistes caeruleus*), great tit (*Parus major*) and chaffinch (*Fringilla coelebs*), are resident and the other, pied flycatcher (*Ficedula hypoleuca*), is a long-distance migrant, wintering in West Africa. All but the chaffinch is hole-nesting. The British Trust for Ornithology (BTO) Nest Record Scheme (NRS) has collated spatially referenced observations of first laying since 1939 (Crick *et al.*, 2003), and for the period up to 2011, this resulted in 38 185 blue tit, 34 455 great tit, 10 167 chaffinch and 18 376 pied flycatcher records (Table S1). Recording has been highly heterogeneous in time and space, with the number of records increasing over time and toward the south of Britain. The only filtering we applied to the data was to discard the records with incorrect latitude and longitude. The data used here are available on request from http://www.bto.org/research-data-services/data-services/data-request-system for the purposes of validating the results of this study.

In most cases, nests are not visited daily and an earliest and latest first lay date is calculated by combining information collected over repeated visits before and after laying. This information includes the (i) date of any previous visit when no eggs were found, (ii) number of eggs, (iii) laying rate, (iv) incubation period, (v) stage of nestling feather development and (vi) nestling developmental rates (Crick *et al.*, 2003). Our phenology measure in analyses is the midpoint between the earliest and latest predicted first egg date.

We used Met Office data on daily mean temperatures for the period 1960–2011 interpolated over the UK at 5 × 5 km resolution (Perry & Hollis, 2005; Perry *et al.*, 2009, www.metoffice.gov.uk/climatechange/science/monitoring/ukcp09/). For 5 km cell centroids, we calculated daily daylengths (in minutes) as the time from sunrise to sunset, based on the equations in the study of Meeus (1991). 5 × 5 km cells were assigned to 150 × 150 km grid cells, which we treated as populations for the purpose of statistical modeling. We chose 150 × 150 km grid cells because these grid cells are reasonably well replicated (e.g., we have blue tit records from 23 grid cells – Table S1) but are at a scale large enough that the between 150 × 150 km grid cell regression of phenology on temperature may approach the asymptotic slope expected as distances become infinite (Appendix S1). For the pied flycatcher, which has a dispersal distance an order of magnitude greater than the other species (Paradis *et al.*, 1998), this condition may not be met, but insufficient information is available to say so with confidence. In the future, we intend to extend these analyses so that we can treat space in a continuous manner. This would avoid the issue of how to discretize space and would also allow us to fully incorporate the effects of spatial autocorrelation within and between grid cells.

Analyses

All models included phenology as a response variable and population (150 km grid cell ID), year, each 25 × 25 km grid cell by year combination and a residual term as random terms. The purpose of including all 25 × 25 km grid cell by year combinations as random effects was to deal with spatiotemporal nonindependence of temperature and phenology observations, the most extreme case being that phenological observations from the same 5 km grid cell and year share the same interpolated average temperatures. First egg date was also regressed on half the difference between the latest and the earliest predicted first egg date, with regression slopes allowed to vary as random effects over records. By doing this, we assume that the actual unobserved first egg dates are normally distributed and lie between the reported earliest and latest first egg date with the same probability for all observations of a species. This probability (i.e., confidence interval) is a function of the variance in regression slopes and is estimated from the data (see Appendix S2).

We considered two null models that included no cues. The first, null, included only the intercept as a fixed effect and the second, year_space, included year (as a continuous term), latitude, longitude and the latitude:longitude interaction as fixed effects. The motivation behind considering these models was

as a baseline against which the performance of models that incorporated cues could be compared. The terms and number of parameters for each model are reported in Table 1.

Remaining models also included year (as a continuous term), latitude and longitude as fixed effects, but also included the action of one or both of photoperiod and temperature as cues (Fig. 1).

For models in which temperature was a cue, we modeled phenology and temperature as a bivariate response to decompose their covariance into various spatial and temporal slopes (see below).

Avtemp_i (Fig. 1a) corresponds to a scenario where phenology responds linearly to mean temperature during a sliding window (Husby *et al.*, 2010; Phillimore *et al.*, 2010). This model assumes that populations in different locations all respond to temperature during the same window. We considered a range of start dates (ordinal dates 1–121, in 2-day increments) and durations for the window (4–120 days, in 2-day increments) and included the end date of the sliding window as an offset in the bivariate model. Here, the phenology intercept can be interpreted as the average lag time between the end of the sliding window and first egg date.

Table 1 Summary of the syntax and number of parameters for each model

Model	Response	Model terms*	k†
null	Univariate	Fixed = intercept Random = grid150, yearF, grid25:year, residual, measurement uncertainty	6
year_space	Univariate	Fixed = intercept, yearC, latitude, longitude, latitude: longitude Random = grid150, yearF, grid25:year, residual, measurement uncertainty	10
photo_only	Univariate	Fixed = intercept Random = grid150, yearF, grid25:year, residual, measurement uncertainty Additional parameter = photoperiod threshold.	7
avtemp_i	Bivariate	Fixed = intercept, yearC, latitude, longitude Random = grid150, yearF, grid25:year, residual, measurement uncertainty Additional parameters = window start date, window duration.	15
avtemp_latvar	Bivariate	Fixed = intercept, yearC, latitude, longitude Random = grid150, yearF, grid25:year, residual, measurement uncertainty Additional parameters = window start date, window duration, latitudinal slopes in window start and duration.	16 if latitudinal slope in window start date or duration is nonzero. 17 if both terms are nonzero.
tempshift_i	Bivariate	Fixed = intercept, yearC, latitude, longitude Random = grid150, yearF, grid25:year, residual, measurement uncertainty Additional parameters = window start date, window duration.	15
phototemp_i	Bivariate	Fixed = intercept, yearC, latitude, longitude Random = grid150, yearF, grid25:year, residual, measurement uncertainty Additional parameters = photoperiod threshold, window duration	15
phototemp_latvar	Bivariate	Fixed = intercept, yearC, latitude, longitude Random = grid150, yearF, grid25:year, residual, measurement uncertainty Additional parameters = photoperiod threshold, window duration, latitudinal gradient in window duration	16

*YearC and YearF denote year as continuous and as a factor, respectively.
†Random terms contribute two parameters to bivariate models, with the exception of the measurement uncertainty term (1 parameter).
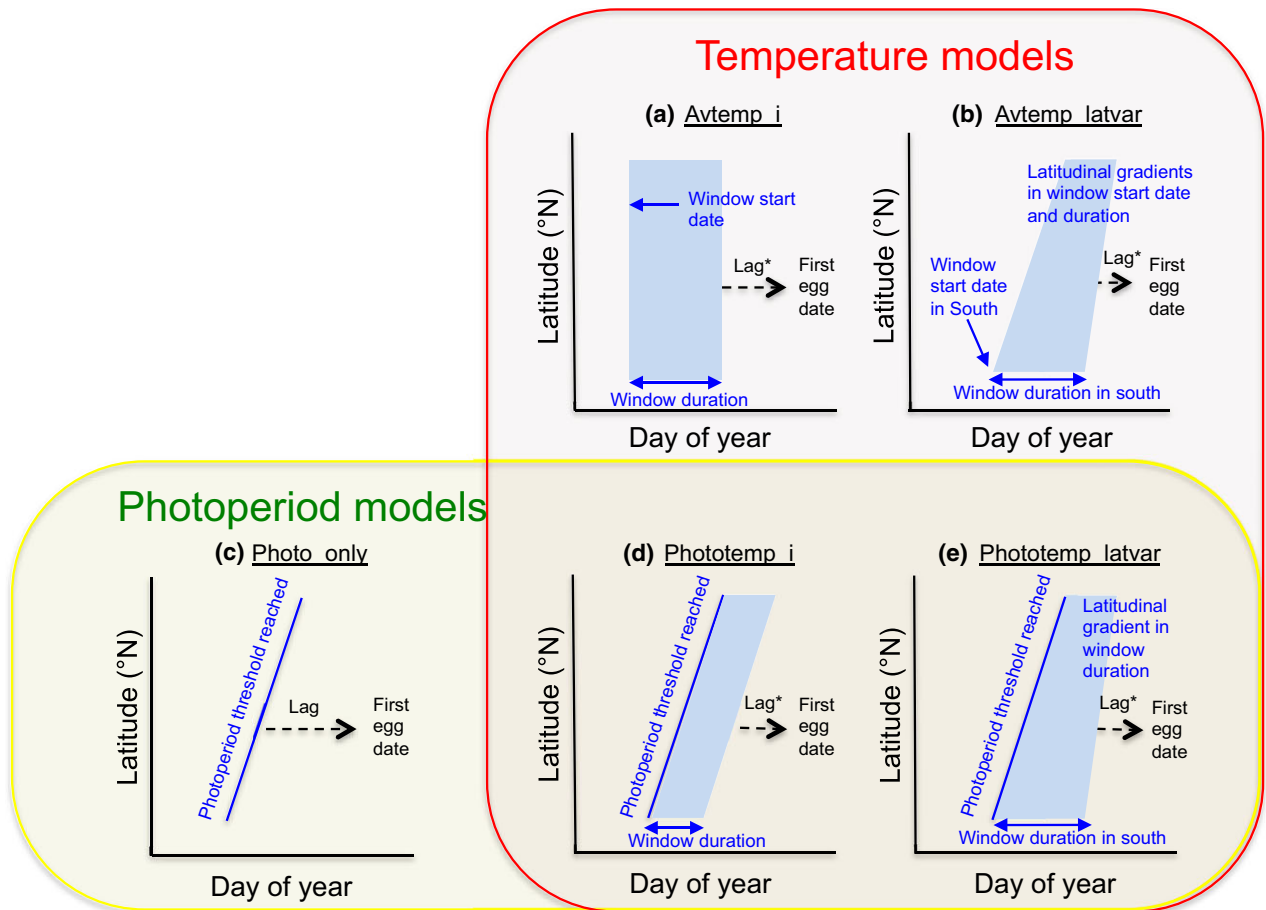
Fig. 1 A schematic of how parameters vary with latitude under models that include (a, b) temperature, (c) photoperiod and (d, e) their combination. Parameters that are optimized via iterative searches (see Table 1) are in blue. Lag* indicates models where the lag duration is a linear response to spatial and temporal variation in the mean (or shift in) temperature during the time window.

Avtemp_latvar (Fig. 1b) models depart from the avtemp_i models in allowing for a latitudinal gradient in the start (slope varied from −5 to 5 days per °North in intervals of 1) and/or duration (slope varied from −5 to 5 days per °North in intervals of 1, with a duration range of 20–100 days at 50°N) of the sliding window over which average temperature affected first egg date. We focused our iterative search of parameter space to start dates in the range 51–121 ordinal days. We were motivated to include this phenomenological model, on the basis of evidence for a shortening and delaying of the sliding window with increasing latitude (Slagsvold, 1976). However, we have no cue-based mechanism on which to predict a latitudinal gradient in the sliding window duration or start date (where the latitudinal gradient in start date departs from the gradient of constant daylength).

The tempshift_i model differed from the avtemp_i model in that phenology is assumed to respond linearly to the *change* in mean temperature (in units of °C day^{-1}) over a sliding window, as estimated using a linear model. We included this model as there has been experimental evidence that lay date in great tits is sensitive to increasing temperature (Schaper *et al.*, 2012).

Photo_only (Fig. 1c) included the day of the year on which a specific daylength was exceeded for each observation as an offset in the model. We iteratively fit models for daylengths spanning the range of 486–876 min in 2-min increments. For this model, the intercept can be interpreted as the average lag time between the threshold being met and first egg date. This model was included for completeness, but we know that it cannot explain interyear variation in phenology.

The phototemp_i (Fig. 1d) model included a photoperiod threshold (from 486 to 876 min in 2-min intervals) to initiate phenological sensitivity to average temperature over a window of specified duration (2–120 days in 2 day intervals). As with the avtemp models, the end of the sliding window was included as an offset on first egg date.

The phototemp_latvar (Fig. 1e) model departed from the phototemp_i model only in that the duration of the sliding window was permitted to vary latitudinally (slope varied from −5 to 5 days per °North in intervals of 1). As with the avtemp_latvar model, we have no mechanistic basis for predicting a latitudinal gradient in the sliding window duration.

We have shown elsewhere that it is possible to separate the contributions of temperature-mediated local adaptation and

plasticity from spatiotemporal data (Phillimore *et al.*, 2010). The premise of this approach is laid out below. In a bivariate mixed model, the slope of phenology (*P*) on temperature (*T*) can be estimated separately for each random effect (*R*), that is, grid cell, year and residual, as

$$\frac{\sigma_{PR,TR}}{\sigma_{TR}^2}$$

In Phillimore *et al.* (2010), we assumed that the temporal (among year) slope of first egg date on the temperature variable was the result of phenotypic plasticity alone. To justify this assumption, we noted that evolutionary change in year *t* is the result of selection occurring in year *t* − 1 and so the temperatures in years *t* − 1 and *t* must be correlated if any of the covariation between temperature and phenology in generation *t* is due to microevolution (Michel *et al.* 2014). Generally, autocorrelation in annual temperatures is weak (for Central England (Parker *et al.*, 1992), annual March–May average temperatures $\rho = 0.45$, reducing to 0.23 after the temporal trend is removed) and so the association between breeding value and temperature across years within populations is expected to be relatively small.

In contrast, the spatial (among grid cell) slope is expected to be the result of plasticity plus local adaptation (Phillimore *et al.*, 2010) because the evolutionary responses at a location may have arisen over many generations. Assuming that all populations share the same plastic response, the difference between the temporal and spatial slopes provides an estimate of the contribution of temperature-mediated local adaptation.

Here, we extend this approach to show that by including latitude, longitude and year (as a continuous variable) as fixed effects in our bivariate model, we are able to estimate (i) the temperature sensitivity of selection on phenology (*B*) over latitude and longitude (Hadfield, in press) and (ii) the sustained phenological response (plasticity plus microevolution) to rising spring temperatures. Estimating these properties is possible due to directional trends in temperature across latitude, longitudes and over time, which facilitates the buildup of an association between temperature and breeding values. We assume that we have identified the sole cue that is causative for phenology (Chevin & Lande, 2015), populations are in migration/selection balance and population density is constant over a species' range. By dividing the coefficient for first egg date on latitude (or longitude) by the coefficient for temperature on latitude (or longitude), we can estimate *B*, the temperature sensitivity of selection (Hadfield, in press). In the absence of confounding variables, the temperature sensitivity of selection is expected to be the same whether estimated across latitudes (B_{lat}) or longitudes (B_{lon}). The difference between this gradient and the temporal slope of phenology on temperature (estimating plasticity, *b*) provides an estimate of the contribution made by clinal local adaptation. $|B - b|$ can be used to estimate the degree to which microevolution will be required in response to rising temperatures. By including latitude and longitude as fixed effects, this means that the slope of phenology on temperature estimated over 150×150 km grid cells corresponds to detrended variation in phenology regressed on detrended variation in the cue. We

anticipate that if local adaptation is imperfect, this slope will be intermediate between *B* and *b*.

We can estimate the sustained phenological response to rising temperature using the same approach but replacing latitude with year. We predict that the slope of the sustained phenological response should be intermediate between *B* and *b* and that a departure from *b* may reveal the contribution of adaptive microevolution during the focal time period. Inclusion of year as a fixed continuous term means that the detrended temporal slope (estimating *b*) will correspond to the phenological response to stochastic, rather than directional, change in temperature (Anderson *et al.*, 2012).

For models where the position and duration of the sliding window varies latitudinally, we assume that any translocated individual would plastically adopt the local sliding window. However, geographic variation in the sliding window might conceivably arise via local adaptation. With the exception of photoperiod-threshold models (Phillimore *et al.*, 2012, 2013), this is the first time that models with latitudinal variation in sliding window start date and duration have been tested in this framework. As we lack a mechanism to explain any latitudinal gradient in the start or duration of the window that is not related to photoperiod (Fig. 1b, e), we also present the results of the best-fitting model where the window is underpinned by a hypothetical mechanism (Fig. 1a, c, d).

We used ASReml-R to obtain restricted maximum likelihood (REML) estimates of model parameters. Rather than relying on the temporal pseudo R^2, as in earlier work (Phillimore *et al.*, 2012), we derived an information criterion. We cannot compare models on the basis of their bivariate likelihood because the temperature data changes from one model to the next. Instead, we calculated the likelihood of phenology (y_1), conditional on temperature (y_2) and REML parameters from a bivariate model ($\hat{\theta}$):

$$L(y_1|y_2, \hat{\theta}) = L(y_1, y_2|\hat{\theta})/L(y_2|\hat{\theta})$$

where the numerator is the likelihood of the bivariate model, and the denominator is the likelihood of a univariate temperature model conditional on the parameters estimated in the bivariate model. Because models differed in their fixed effects, restricted likelihoods are not comparable, and so standard likelihoods were used. It should be noted that the estimates ($\hat{\theta}$) are those that maximize the restricted likelihood and are therefore not the maximum likelihood estimates, although the differences should be small. For each model, we calculated the Akaike information criteria (AIC) and used AIC weights (Burnham & Anderson, 2004) to compare the relative support received by our eight classes of model. In calculating the number of parameters (*k*) for each model, we included the parameters that we used to define the photoperiod threshold and sliding window. For bivariate models, each random term requires estimation of three covariance parameters, but for the conditional model we consider each random term as contributing only two (the regression of phenology on temperature $\beta_{PT} = \sigma_{PR,TR}/\sigma_{TR}^2$ and the unexplained variation in phenology $\sigma_{PR}^2(1 - \beta_{PT}^2)$). The extent to which the multiple testing inherent in sliding window analyses influences the interpretation of their results has not been studied and may introduce a bias toward selecting more complex models

(Michael Morrissey pers. comm). However, the very high autocorrelation between temperatures on consecutive days will serve to reduce the bias.

Bayesian parameter estimation

For each species, we selected two focal models that included temperature as a cue: (i) the lowest AIC avtemp_latvar model, which is our most complex phenomenological model and (ii) the lowest AIC model phototemp_i model, which is our most complex model for which we have a mechanism for the timing of the start of the sliding window. We refitted each focal model in a Bayesian mixed model environment using MCMCglmm (Hadfield, 2010), which allowed us to calculate the 95% credible intervals on derived slopes and the differences between them. Priors for the (co)variance components followed the inverse Wishart distribution with $V = I$ and $v = 0.002$.

All analyses were conducted in the R environment (R Development Core Team, 2011).

Results

Spatiotemporal patterns

Each nest record comes with a predicted earliest and latest first egg date. Assuming that measurement uncertainty follows a normal distribution, we estimate that this range represents approximately the 81% confidence interval for pied flycatcher and chaffinch and the 94% confidence interval for blue and great tits. For all species, the largest variance component is the residual term, that is, the variance within a 25×25 km grid square in a single year (Fig 2a). For 25×25 km grid cells and years with at least one observation, the median number of records is two for blue tit, great tit and chaffinch and four for pied flycatcher. The estimated residual variances imply that 95% of first lay dates within a single site and year will lie within a 26-, 32- and 28-day periods for blue tits, great tits and pied flycatchers, while for chaffinch this period is 52 days. The greater variance in chaffinch first lay date may be due to the presence of second breeding attempts caused by higher levels of nest-predation in open nesters (Bennett & Owens, 2002). For blue tit, great tit and pied flycatcher, the variance in first egg dates among years and among 150×150 km grid cells is broadly comparable, with considerably less variance distributed among 25×25 km:year. Chaffinches depart from this pattern in that the 25×25 km:year variance component exceeds the variance among 150×150 km grid cells.

First egg dates have advanced significantly over the past 50 years for all species (Fig 2b), with great tits exhibiting the most rapid advance in first egg date of 0.36 ± 0.05 days per year (Table S2). The first egg dates of all four species show significant geographic trends
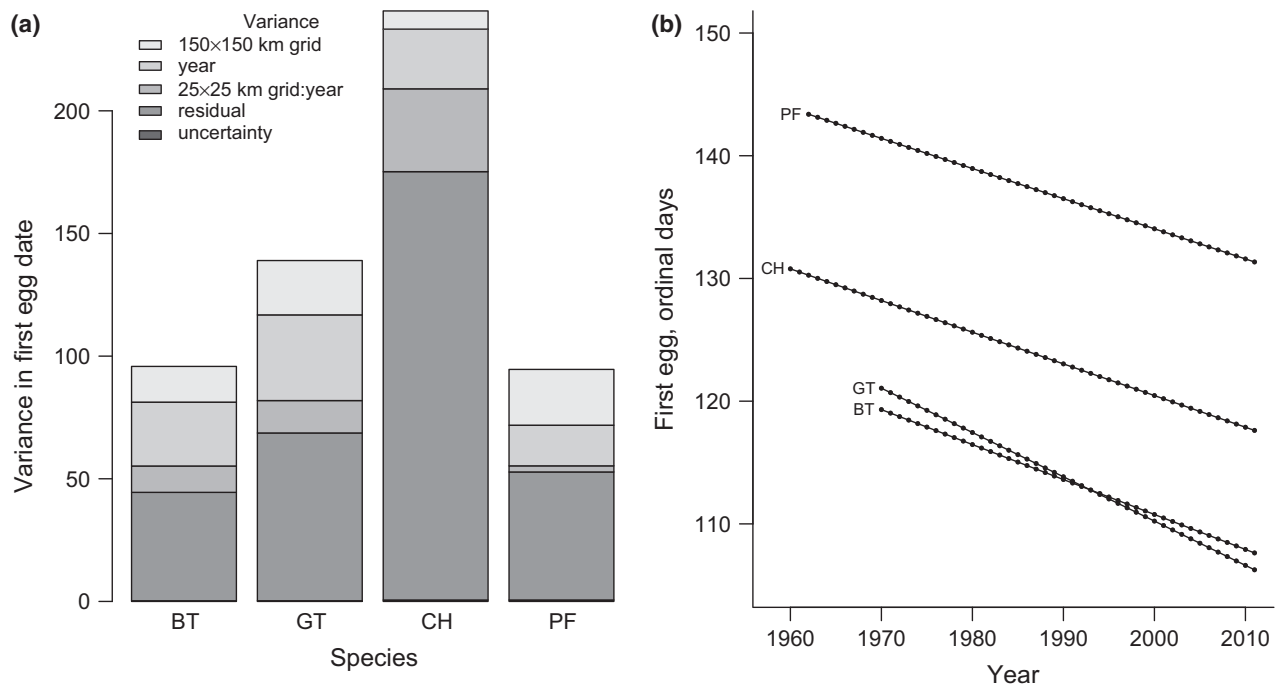

Fig. 2 (a) Variance components from the null model. (b) Model predictions for linear change in first egg dates over time from the year_space model for latitude = 51.4791 and longitude = 0. Dots represent years with data. Species codes are BT, blue tit; GT, great tit; CH, chaffinch and PF, pied flycatcher.
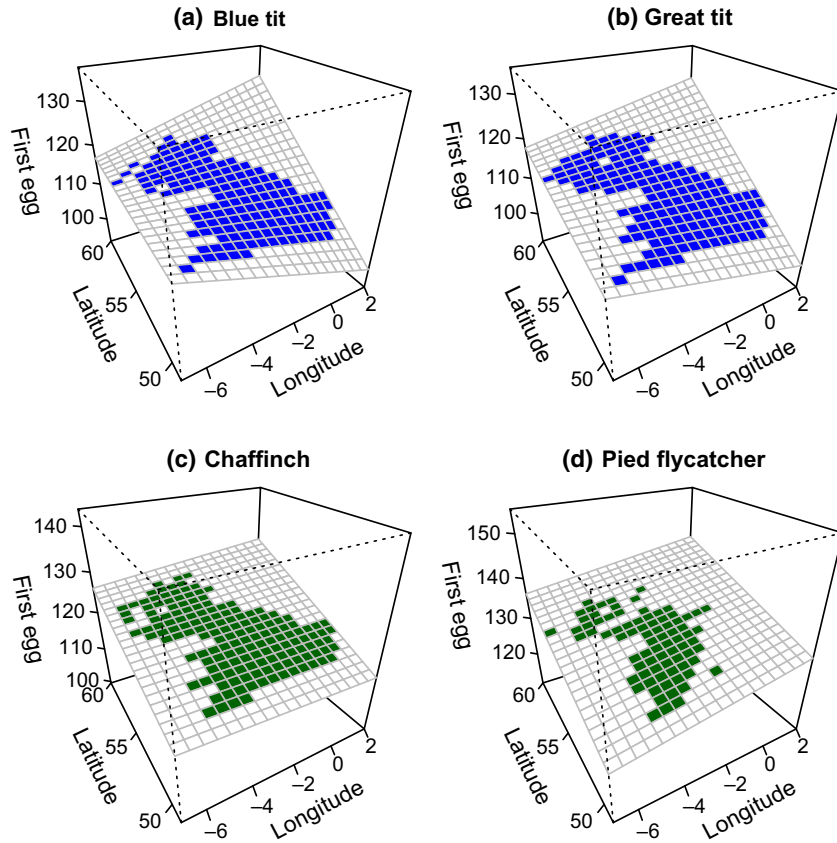
Fig. 3 A 3D plot of the latitudinal and longitudinal trends in first egg dates. Filled 25 × 25 km grid cells correspond to geographic locations with records. Slope estimates obtained from the year_space null model (see methods for details) fixing year = 2011. Blue squares indicate species for which the latitude:longitude interaction was significant. Green squares indicate species for which deletion of latitude and longitude led to a significant likelihood ratio test.
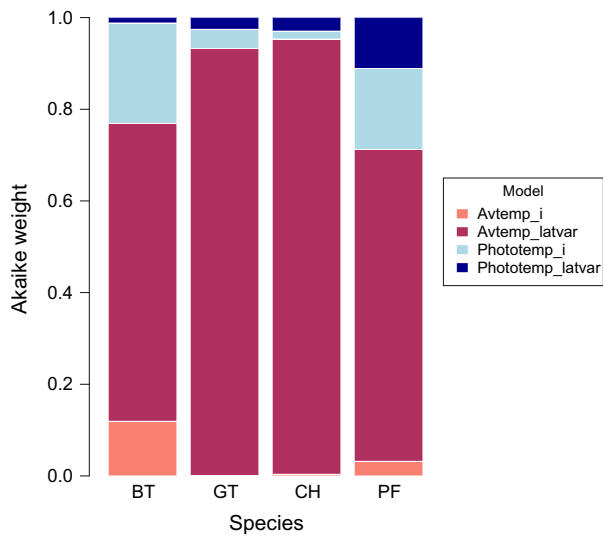

Fig. 4 Akaike weights comparing among all models. Null, year_space, photo_only and tempshift_i models had Akaike weights <0.001.

(Fig 3), with first egg date delayed as latitude increases. For blue and great tits, first egg dates are earliest in the southeast and there is a significant interaction between latitude and longitude, such that the latitudinal gradient in first egg date is steeper in the east than the west of Britain.

Environmental predictors

Models that included average temperature as a predictor receive strong AIC support (Fig. 4, Table S2). Sliding windows identify a period of temperature sensitivity spanning 1–3 months in the period March–May and overlapping with the distribution of first egg dates. For each species, the avtemp_latvar model (Fig. 1b) returns the lowest AIC, identifying a trend for the start of the sliding window to be delayed further north, with the latitudinal gradient steepest in great tit and pied flycatcher (Fig. 5, Table S2). For great tit and pied flycatcher, we also find that the sliding window becomes shorter in duration further north (ΔAIC vs.
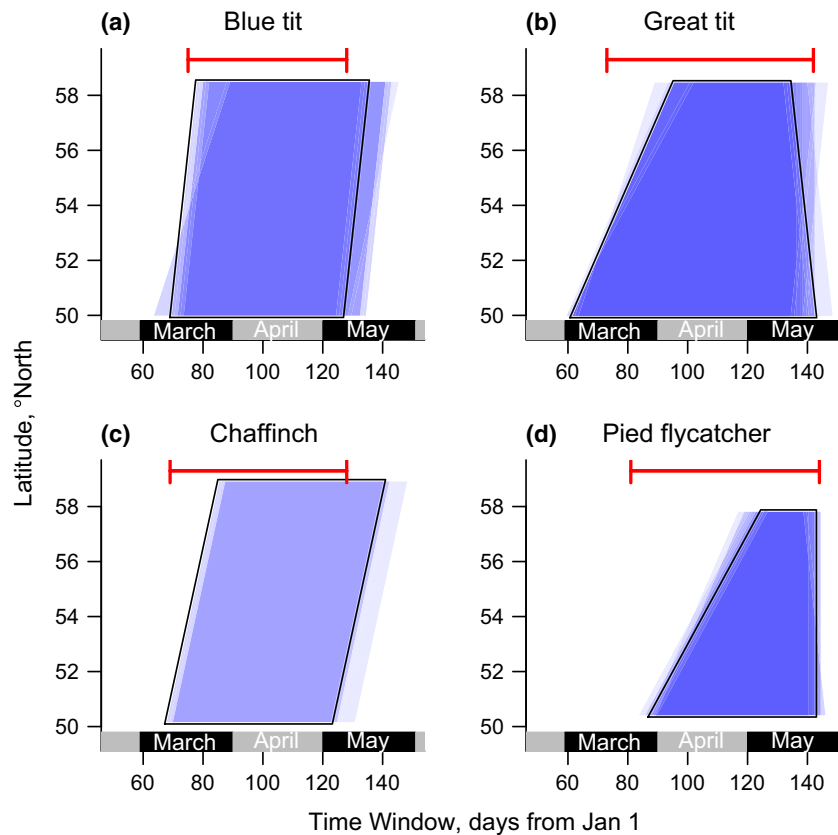
Fig. 5 Latitudinal trends in the avtemp_latvar sliding windows over which mean temperature best predicts spatiotemporal variation in first egg dates. The sliding window identified by the lowest Akaike information criteria (AIC) model is outlined in black, with transparent blue polygons reporting alternative models within 2 ΔAIC of the best model. The lowest AIC avtemp_i window is shown in red at the top of each plot.

best model with no latitudinal gradient in sliding window duration is 6.2 in great tit and 2.7 in pied flycatcher).

While models combining a photoperiod threshold and sliding window approach (1d and 1e) are not the best performing as measured by AIC, for each species they are within ~8ΔAIC of the best model (Table S2, Fig. 4). The best photoperiod thresholds correspond to times immediately prior to the equinox for blue tit, great tit and chaffinch and immediately after for pied flycatcher (Table S3). This means that this model generates a very shallow latitudinal gradient in the sliding window start date. For great tit and chaffinch, the ΔAIC between the best model (avtemp_latvar) and the best photo-temp model exceeds 6, and the difference between models appears to support a greater delay in the initiation of temperature sensitivity with increasing latitude than expected if initiation was due to a single photoperiod threshold (Table S2). We find no evidence that populations are responding to a shift in temperature during a sliding window (ΔAIC > 69, Table S2).

Plasticity, microevolution and the temperature sensitivity of selection

Detrended temporal slopes, estimated as the slope of interannual variation in phenological lag regressed on average temperature, and which we assume are attributable to plasticity (*b*), are significantly negative for all species (Fig. 6a, b, Table S3). This response varies from an advance of approximately 2.3 days per °C in pied flycatcher up to 4.8 days days per °C in great tits. These slopes are relatively insensitive to the model used to define the sliding window (Table S2). Under the phototemp_i model that best captures the effects of known cues, we find that when we estimate the temporal slopes separately for each 150 × 150 km grid cell, the slopes are similar, remarkably so, when we consider that each slope is estimated with error (Fig. S1). This implies that our model assumption of constant plasticity is not violated.

Our estimates of the temperature sensitivity of selection (B_{lat}) are highly sensitive to the latitudinal gradient in the end date of the sliding window that defines
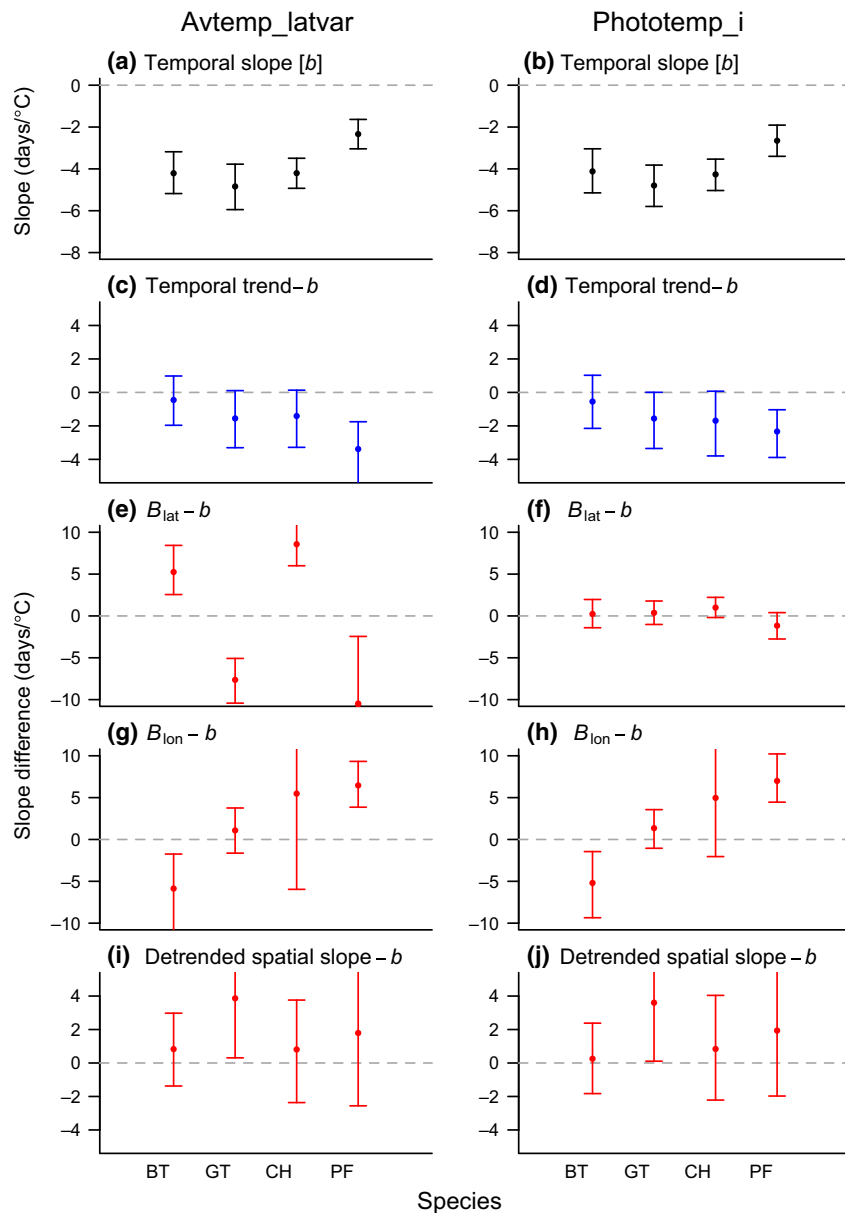
Fig. 6 Posterior medians and 95% HPDs for (a, b) slopes and (c–j) slope differences for the best avtemp_latvar (a, c, e, g, i) and phototemp_i models (b, d, f, h, j).

thermal sensitivity (Table S2). This is because any latitudinal gradient in the end of the sliding window will affect both the latitudinal gradient in phenological lag time and the average temperature calculated over the time window. Under the lowest AIC models, these slopes are significantly negative in great tit and pied flycatcher, and positive in chaffinch (Table S3a) and b is unable to track the optimum for all species (Fig. 6e). However, if we consider the lowest AIC phototemp_i model, B_{lat} is estimated to be statistically indistinguishable in magnitude from our estimate of b for all four species (Fig. 6f).

Estimates of the longitudinal temperature sensitivity of selection (B_{lon}) generally have wider credible intervals than the estimates of B_{lat}, which can be attributed to phenology and temperature varying less over longitude than latitude. B_{lon} estimates are inconsistent across species but largely consistent between models (Table S3). B_{lon} is significantly negative for blue and great tit and positive for pied flycatcher, and the gradient departs significantly from our estimate of plasticity for blue tit and pied flycatcher (Fig. 6g, h). Estimates of B_{lon} differ substantially from estimates of B_{lat} across all species for the avtemp_latvar model and across blue tit

and pied flycatcher for the phototemp_i model (Table S3), which is hard to reconcile with theory.

Under the best avtemp_latvar and phototemp_i models, the relationship between first egg date and the detrended spatial variation in average temperatures is negative for all species, but only significantly so for blue tit and chaffinch (Table S3). The detrended spatial slope is shallower than b for all four species, although in no case is the slope difference significant (Fig. 6i, j), which is consistent with the null expectation of no local adaptation to temperatures that vary nonclinally in space.

The change in phenology in response to directional change in temperature over the past half century is significantly negative for all species (Table S3) and steeper than our estimates of plasticity (Fig. 6c, d). For pied flycatcher, the slope difference is significant and, in agreement with our theory-derived expectations, the slope difference is similar in sign and magnitude to the difference between B and b, consistent with microevolution compensating for imperfect plasticity (Table S3).

Discussion

Blue tit, great tit, chaffinch and pied flycatcher lay dates are sensitive to spring temperature variation across time and space, with warmer temperatures eliciting earlier phenology. We find no compelling evidence that temperature-mediated local adaptation contributes to among population differences in their phenology. Taking the parameter estimates from the best performing phototemp_i model (the most complex model for which we can justify the parameterization of the sliding window solely on the basis of known cues), we find that phenological plasticity tracks temperature-mediated variation in the optimum across latitudes. Substituting space for time, this implies that all four species possess adaptive plasticity (b) that should allow them to cope with even rapid changes in temperature, as $|B - b|$ is small (Chevin et al., 2010). This result is also consistent with longitudinal studies that find lay date plasticity partially or completely tracks year-to-year variation in the optimum (Charmantier et al., 2008; Both et al., 2009) and may continue to do so into the future (Gienapp et al., 2013; Vedder et al., 2013). However, we caution that our inferences regarding B_{lat} are highly sensitive to latitudinal trends in the end date of the sliding window and that estimates of $B_{lat} \neq B_{lon}$ in some cases.

Our finding of no temperature-mediated local adaptation of passerine lay dates across Britain is on the face of it surprising, given estimates of heritability and selection on lay date and reports that great tit body mass and clutch size show local adaptation on a much finer spatial scale (Postma & Van Noordwijk, 2005).

However, if plasticity is not costly and enables populations to track temperature-mediated variation in optimum phenology over space, as appears to be the case with lay date, then selection for local adaptation may be weak or absent (Via & Lande, 1985). In addition, populations are also likely to be connected by high levels of gene flow (Van Bers et al., 2012), which will reduce the efficiency of local adaptation (Lenormand, 2002) to track any local deviations from the latitudinal/longitudinal trend in temperature. While we find no significant evidence for local adaptation in chaffinch, we note that for this species the spatial slopes are consistent with countergradient variation, being shallower than the plastic response (Conover & Schulz, 1995). Countergradient variation may be an adaptation to a shortening of the growing season with increasing latitude (Conover & Schulz, 1995) and has been reported for the spring phenology of several UK plant (C.J. Tansey, J.D. Hadfield, A.B. Phillimore, unpublished data) and butterfly (Roy et al., 2015) species.

One promising avenue that this work presents is the estimation of B from existing observational data. For great tit, our estimates of B under the phototemp_i model ($B_{lat} = -4.41$, $B_{lon} = -3.45$) are comparable to estimates for the Wytham Woods ($= -5.30$, Vedder et al., 2013) and Hoge Veluwe populations ($= -5.01$, Chevin et al., 2015). Given the importance of this parameter in climate change projections (Chevin et al., 2010), here we briefly consider statistical issues and theoretical limitations. First, this approach is correlational and we may not have correctly identified the environmental variable(s) that the populations are adapting to and/or the environmental variable(s) that the populations are responding plastically to (Chevin & Lande, 2015). In both cases, the results will not be robust if the spatial correlation structure between our putative cue and the true driving environmental variables differs from their temporal correlation structure. However, if the spatial and temporal correlation structures are similar, then all terms (e.g., B, b) will be multiplied by the regression of the putative cue on the true environmental variables and their relative magnitudes should remain comparable (Michel et al. 2014, Hadfield, in press). In the case where the putative cue is actually the environmental variable, the populations are adapting to, but this differs from the environmental variable that determines the plastic response, then b is best interpreted as the degree to which plasticity tracks the environment of selection. Under these circumstances, b is expected to be shallower than B (Tufto 2015) even when plasticity is cost-free. The fact that we do not observe this gives us some hope that the environment of selection and the environment that determines plasticity are strongly correlated.

Our estimates of B_{lon}, which in some cases depart substantially from the plasticity and B_{lat} slopes, may reflect the action of additional cue(s) not included in our analysis (Chevin & Lande, 2015). Recent work on a great tit population demonstrates how the optimum lay date correlates with interannual variation in the timing of peak caterpillar availability (Chevin *et al.*, 2015). If forest types vary geographically, this could affect the mean and variance of prey availability in different areas (Burger *et al.*, 2012), which may lead to the relationship between temperature and the food-peak, and thereby B, to vary spatially. We suggest that this may not present a major problem for our analyses, as deciduous woodland is distributed across Britain. However, over larger spatial scales, where there are major habitat transitions, our approach for estimating B may not be appropriate.

Even in the absence of third variables, it is possible that rather than plasticity tracking the optimum, we may in fact have underestimated the true gradient in the optimum, either due to populations having not reached migration/selection equilibrium or due to changes in population density with latitude or longitude. Differences in population density generate asymmetric gene flow from high-density to low-density areas, resulting in greater maladaptation in low-density areas (Pease *et al.*, 1989). If population density is greatest in the center of a species range and declines toward the periphery, then the latitudinal/longitudinal trend will be shallower than the optimum, and consequently B will be underestimated (Pease *et al.*, 1989; García-Ramos & Kirkpatrick, 1997). While relative abundance of chaffinches is quite consistent across Britain, abundance declines in northern regions (Scotland) for the remaining species (Balmer *et al.*, 2013). However, given that most of the geographic variation in phenology appears to be due to plasticity rather than local adaptation, the effects of asymmetric dispersal on the spatial distribution of phenotypes is likely to be small.

Numerous studies on wild systems have sought to estimate the contribution of adaptive microevolution in response to temporal changes in the environment (e.g., Sheldon *et al.*, 2003). Despite such efforts, in very few cases is the evidence for microevolution compelling (Hansen *et al.*, 2012; Merilä, 2012). We find some evidence consistent with a response to selection for earlier lay date in pied flycatcher over the last 50 years, in that lay date has advanced by significantly more in response to the increase in spring temperature than expected given our estimate of plasticity. Moreover, the difference between these slopes (posterior median = −2.3) is of similar sign and magnitude to the difference between our best estimate of the temperature sensitivity of selection and the degree to which this is tracked by plasticity

(posterior median of $B_{lat} - b = -1.1$), which should determine the strength and direction of selection. Consistent with this interpretation, the pied flycatcher is the least plastic of the focal species, meaning that it may be least able to track the phenology of woodland invertebrates via plasticity, and for a Dutch population selection has been for earlier laying in most years (Visser *et al.*, 2015). However, we suggest caution in interpreting our finding as rare evidence for microevolution in response to climate change as (i) our inference is reliant on our inferred cue being closely correlated with the true cue and (ii) our test relies on the estimation of three separate slopes, making it especially vulnerable to the action of one or more third variables. For the remaining species, while the magnitude of the response to directional temperature change also exceeds b they are not significant. It may be that there has been no adaptive response in these species or that plasticity tracks the optimum sufficiently well for selection to be weak and the response hard to detect (Merilä, 2012).

Our analyses provide strong evidence that mean spring temperatures during specific sliding windows are negatively correlated with lay date, as is well established (Crick & Sparks, 1999; Husby *et al.*, 2010). Consistent with earlier work on great tits (Slagsvold, 1976), we find a delay in the initiation of thermal sensitivity as latitude increases, which begs the question what are the conditions or cues that initiate the period of temperature sensitivity? Although models combining photoperiod and average temperature are not the very best performing in terms of AIC, we suggest that these models are the most biologically defensible given current understanding of phenological cues (Caro *et al.*, 2013). Alternatively, the latitudinal gradients in the sliding window start dates and durations (avtemp_latvar model) may capture real geographic variation in cues. If this were the case, then the B_{lat} estimates obtained for this model may indicate a violation of one or more of the assumptions that we make in estimating B. In contrast to Schaper *et al.* (2012), we find no evidence that a gradient in temperature, as opposed to the average temperature, best predicts variation in first egg dates.

Our estimates of b, B and the detrended spatial slope are all predicated on lay date (and the optimum) being determined by an environmental cue that has been correctly identified. While the sliding window approach we adopted identifies the period over which the temperature–phenology correlation is strongest, unresolved issues remain with the notion that average temperature is the direct cue. First, we find, as many others have before (Perrins, 1965; Gienapp *et al.*, 2005), that the most predictive window overlaps substantially with the period of laying (see the high frequency of negative lag times in Fig. S1), implying, nonsensically,

that birds cue off temperatures after laying to determine when to lay. In defense of the sliding window approach, the effects that are identified may correlate closely with the true mechanism, as has been shown for growing degree day models in plants (Phillimore *et al.*, 2013; Roberts *et al.*, 2015). Second, with respect to the migratory pied flycatcher, temperature sensitivity (as identified by the phototemp_i model) is initiated around March 21, whereas first arrival dates on breeding grounds in forests in southwest England and northeast Wales average approximately 20 days later (M. Burgess and B. Harris pers. comm). This mirrors earlier findings in relation to migratory collared flycatchers (Löhrl, 1957). The period of temperature sensitivity prior to return to the breeding grounds is possibly due to spatial autocorrelation between temperatures on migration and breeding grounds (Ockendon *et al.*, 2013; Finch *et al.*, 2014). Alternatively, the effect of temperature on lay date may be indirect, via its effect on the phenology and growth of invertebrate prey, in turn acting as a constraint or cue for laying (Perrins, 1965).

In this study, we show how spatiotemporal phenological observations can be used to identify cues and estimate two parameters, plasticity and the environmental sensitivity of the optimum, that are key to projecting population responses under a climate change (Chevin *et al.*, 2010). We find that plasticity may track temperature-mediated variation in optimum lay date for the four focal passerines, which, if true, suggests that populations of these species are well positioned to persist in the face of rising spring temperatures.

Acknowledgements

References

Anderson JT, Inouye DW, Mckinney AM, Colautti RI, Mitchell-Olds T (2012) Phenotypic plasticity and adaptive evolution contribute to advancing flowering phenology in response to climate change. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **279**, 3843–3852.

Balmer DE, Gillings S, Caffrey B, Swann R, Downie I, Fuller R (2013) *Bird Atlas 2007–11: The Breeding and Wintering Birds of Britain and Ireland*. BTO, Thetford.

Bennett PM, Owens IPF (2002) *Evolutionary Ecology of Birds: Life Histories, Mating Systems and Extinction*. Oxford University Press, Oxford, UK.

Blondel J, Perret P, Maistre M (1990) On the genetical basis of the laying-date in an island population of blue tits. *Journal of Evolutionary Biology*, **3**, 469–475.

Blondel J, Dias PC, Perret P, Maistre M, Lambrechts MM (1999) Selection-based biodiversity at a small spatial scale in a low-dispersing insular bird. *Science*, **285**, 1399–1402.

Both C, Van Asch M, Bijlsma RG, Van Den Burg AB, Visser ME (2009) Climate change and unequal phenological changes across four trophic levels: constraints or adaptations? *Journal of Animal Ecology*, **78**, 73–83.

Burger C, Belskii E, Eeva T *et al.* (2012) Climate change, breeding date and nestling diet: how temperature differentially affects seasonal changes in pied flycatcher diet depending on habitat variation. *Journal of Animal Ecology*, **81**, 926–936.

Burnham KP, Anderson DR (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods and Research*, **33**, 261–304.

Caro SP, Schaper SV, Hut RA, Ball GF, Visser ME (2013) The case of the missing mechanism: how does temperature influence seasonal timing in endotherms? *PLoS Biology*, **11**, e1001517.

Charmantier A, Gienapp P (2013) Climate change and timing of avian breeding and migration: evolutionary versus plastic changes. *Evolutionary Applications*, **7**, 15–28.

Charmantier A, Mccleery RH, Cole LR, Perrins C, Kruuk LEB, Sheldon BC (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, **320**, 800–803.

Chevin LM, Lande R (2015) Evolution of environmental cues for phenotypic plasticity. *Evolution*, **69**, 2767–2775.

Chevin L-M, Mace GM, Lande R (2010) Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS Biology*, **8**, e1000357.

Chevin LM, Visser ME, Tufto J (2015) Estimating the variation, autocorrelation, and environmental sensitivity of phenotypic selection. *Evolution*, **69**, 2319–2332.

Conover DO, Schultz ET (1995) Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution*, **10**, 248–252.

Crick HQP, Sparks TH (1999) Climate change related to egg-laying trends. *Nature*, **399**, 423–424.

Crick HQP, Baillie SR, Leech DI (2003) The UK Nest Record Scheme: its value for science and conservation: capsule A review of its methodology, analytical procedures and uses. *Bird Study*, **50**, 254–270.

Dawson A, King VM, Bentley GE, Ball GF (2001) Photoperiodic control of seasonality in birds. *Journal of Biological Rhythms*, **16**, 365–380.

Dunne JA, Saleska SR, Fischer ML, Harte J (2004) Integrating experimental and gradient methods in ecological climate change research. *Ecology*, **85**, 904–916.

Finch T, Pearce-Higgins J, Leech DI, Evans K (2014) Carry-over effects from passage regions are more important than breeding climate in determining the breeding phenology and performance of three avian migrants of conservation concern. *Biodiversity and Conservation*, **23**, 2427–2444.

García-Ramos G, Kirkpatrick M (1997) Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, **51**, 21–28.

Gienapp P, Hemerik L, Visser ME (2005) A new statistical tool to predict phenology under climate change scenarios. *Global Change Biology*, **11**, 600–606.

Gienapp P, Teplitsky C, Alho JS, Mills JA, Merila J (2008) Climate change and evolution: disentangling environmental and genetic responses. *Molecular Ecology*, **17**, 167–178.

Gienapp P, Väisänen RA, Brommer JE (2010) Latitudinal variation in breeding time reaction norms in a passerine bird. *Journal of Animal Ecology*, **79**, 836–842.

Gienapp P, Lof M, Reed TE, Mcnamara J, Verhulst S, Visser ME (2013) Predicting demographically-sustainable rates of adaptation: can great tit breeding time keep pace with climate change? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **368**, 20120289.

Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, **33**, 1–22.

Hadfield JD (in press) The spatial scale of local adaptation in a stochastic environment. *Ecology Letters*.

Hansen MM, Olivieri I, Waller DM, Nielsen EE, Group TGW (2012) Monitoring adaptive genetic responses to environmental change. *Molecular Ecology*, **21**, 1311–1329.

Husby A, Nussey DH, Visser ME, Wilson AJ, Sheldon BC, Kruuk LEB (2010) Contrasting patterns of phenotypic plasticity in reproductive traits in two great tit (*Parus major*) populations. *Evolution*, **64**, 2221–2237.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Lambrechts MM, Perret P, Blondel J (1996) Adaptive differences in the timing of egg laying between different populations of birds result from variation in photoresponsiveness. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **263**, 19–22.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.

Löhrl H (1957) Populationsökologische Untersuchungen beim Halsbandschnäpper (*Ficedula albicollis*). *Bonner Zoologische Beiträge*, **2**, 130–177.

Meeus JH (1991) *Astronomical Algorithms*. Willmann-Bell, Richmond, VA.

Merilä J (2012) Evolution in response to climate change: in pursuit of the missing evidence. *BioEssays*, **34**, 811–818.

Michel MJ, Chevin LM, Knouft JH (2014) Evolution of phenotype–environment associations by genetic responses to selection and phenotypic plasticity in a temporally autocorrelated environment. *Evolution*, **68**, 1374–1384.

Ockendon N, Leech D, Pearce-Higgins JW (2013) Climatic effects on breeding grounds are more important drivers of breeding phenology in migrant birds than carryover effects from wintering grounds. *Biology Letters*, **9**, 20130669.

Paradis E, Baillie SR, Sutherland WJ, Gregory RD (1998) Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, **67**, 518–536.

Parker DE, Legg TP, Folland CK (1992) A new daily Central England Temperature Series, 1772–1991. *International Journal of Climatology*, **12**, 317–342.

Pease CM, Lande R, Bull J (1989) A model of population growth, dispersal and evolution in a changing environment. *Ecology*, **70**, 1657–1664.

Perrins CM (1965) Population fluctuations and clutch-size in the Great Tit, *Parus major* L. *Journal of Animal Ecology*, **34**, 601–647.

Perry M, Hollis D (2005) The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology*, **25**, 1041–1054.

Perry M, Hollis D, Elms M (2009) Climate Memorandum No 24: The Generation of Daily Gridded Datasets of Temperature and Rainfall for the UK. National Climate Information Centre, Met Office.

Phillimore AB, Hadfield JD, Jones OR, Smithers RJ (2010) Differences in spawning date between populations of common frog reveal local adaptation. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 8292–8297.

Phillimore AB, Stålhandske S, Smithers RJ, Bernard R (2012) Dissecting the contributions of plasticity and local adaptation to the phenology of a butterfly and its host plants. *The American Naturalist*, **180**, 655–670.

Phillimore AB, Proios K, O'mahony N, Bernard R, Lord AM, Atkinson S, Smithers RJ (2013) Inferring local processes from macro-scale phenological pattern: a comparison of two methods. *Journal of Ecology*, **101**, 774–783.

Pickett ST (1989) Space-for-time substitution as an alternative to long-term studies. In: *Long-Term Studies in Ecology* (ed. Likens GE), pp. 110–135. Springer, New York, NY.

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**, 1041–1051.

Postma E, Van Noordwijk AJ (2005) Gene flow maintains a large genetic difference in clutch size at a small spatial scale. *Nature*, **433**, 65–68.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org (accessed 14 August 2015).

Robb GN, Mcdonald RA, Chamberlain DE, Bearhop S (2008) Food for thought: supplementary feeding as a driver of ecological change in avian populations. *Frontiers in Ecology and the Environment*, **6**, 476–484.

Roberts AMI, Tansey C, Smithers RJ, Phillimore AB (2015) Predicting a change in the order of spring phenology in temperate forests. *Global Change Biology*, **7**, 2603–2611.

Roy DB, Oliver TH, Botham MS *et al.* (2015) Similarities in butterfly emergence dates among populations suggest local adaptation to climate. *Global Change Biology*, **21**, 3313–3322.

Schaper SV, Dawson A, Sharp PJ, Gienapp P, Caro SP, Visser ME (2012) Increasing temperature, not mean temperature, is a cue for avian timing of reproduction. *The American Naturalist*, **179**, E55–E69.

Sheldon B, Kruuk L, Merila J (2003) Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution*, **57**, 406–420.

Slagsvold T (1976) Annual and geographical variation in the time of breeding of the great tit *Parus major* and the pied flycatcher *Ficedula hypoleuca* in relation to environmental phenology and spring temperature. *Ornis Scandinavica*, **7**, 127–145.

Slatkin M (1978) Spatial patterns in distributions of polygenic characters. *Journal of Theoretical Biology*, **70**, 213–228.

Tufto J (2015) Genetic evolution, plasticity, and bet-hedging as adaptive responses to temporally autocorrelated fluctuating selection: A quantitative genetic model. *Evolution*, **69**, 2034–2049.

Van Bers NE, Santure AW, Van Oers K *et al.* (2012) The design and cross-population application of a genome-wide SNP chip for the great tit Parus major. *Molecular ecology resources*, **12**, 753–770.

Vedder O, Bouwhuis S, Sheldon BC (2013) Quantitative assessment of the importance of phenotypic plasticity in adaptation to climate change in wild bird populations. *PLoS Biology*, **11**, e1001605.

Verhulst S, Nilsson J-Å (2008) The timing of birds' breeding seasons: a review of experiments that manipulated timing of breeding. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **363**, 399–410.

Via S, Lande R (1985) Genotype–environment interaction and the evolution of phenotypic plasticity. *Evolution*, **39**, 505–522.

Visser ME, Both C, Lambrechts MM (2004) Global climate change leads to mistimed avian reproduction. *Advances in Ecological Research*, **35**, 89–110.

Visser ME, Holleman LJM, Caro SP (2009) Temperature has a causal effect on avian timing of reproduction. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **276**, 2323–2331.

Visser ME, 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**, e1002120.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. The temporal slopes (blue lines) estimated within 150×150 km grid cells under the lowest AIC phototemp_i model for each species. The bivariate model was as described in the methods, except that latitude, longitude and year were removed as fixed effects and 150×150 km grid cell was removed as a random effect. The model was applied to each grid cell in turn. Red points correspond to the grid cell averages. We only include grid cells with >50 nest record observations and >10 years of replication.

Table S1. Summary of the temporal and geographic replication of nest records for each species.

Table S2. Model support and ASReml parameter estimates for the lowest AIC model for each model type.

Table S3. Posterior medians (95% highest posterior densities) for parameters of interest from the lowest AIC (a) avtemp_latvar and (b) phototemp_i model.

Appendix S1. Grid cell size and local adaptation.

Appendix S2. Interval estimation.