Running head: Fitting phenological curves with GLMMs

Title: Fitting phenological curves with Generalized Linear Mixed Models (GLMMs)

Authors:

Collin Edwards[1] (corresponding author), Department of Biology, Tufts University, Medford MA 02155 USA

Elizabeth E. Crone[2], Department of Biology, Tufts University, Medford MA 02155 USA

Abstract

Understanding organismal phenology has been an emerging interest in ecology, in part because phenological shifts are one of the most conspicuous signs of climate change. While we are seeing increased collection of phenological data and creative use of historical data sets, existing statistical tools to measure phenology are generally either limited (e.g., first day of observation, which has problematic biases) or are challenging to implement (often requiring custom coding, or enough data to fit many parameters). We present a method to fit phenological data with Gaussian curves using linear models, and show how robust phenological metrics can be obtained using standard linear regression tools. We then apply this method to eight years of Baltimore checkerspot data using generalized linear mixed models (GLMMs). This case study illustrates the ability of years with extensive data to inform years with less data and shows that butterfly flight activity is somewhat earlier in warmer years. We believe our new method fills a convenient midpoint between ad hoc measures and custom-coded models.

Key words: *phenology; climate change; peak abundance; activity period; first emergence; Gaussian curve; Euphydryas phaeton, growing degree days, generalized linear mixed model*

[1] collin.edwards@tufts.edu
[2] elizabeth.crone@tufts.edu

Introduction

Ecologists have long been interested in understanding phenology, the timing of life history events. In the face of climate change, there has been increased interest in quantifying and analyzing phenological trends (e.g., Inouye et al. 2019, Chmura et al. 2019). Shifts in phenology are one of the most well-documented signs of biological responses to climate change, with evidence for changes in timing of key life history events such as flowering of plants (Cleland 2007, Gordo 2009), flight activity of insects (Roy and Sparks 2000, Forister and Shapiro 2003), and passage of migratory birds (Marra et al. 2005, van Buskirk et al. 2009, Horton et al. 2020). Nonetheless, in spite of widespread interest in phenology, analyses of phenological variables from monitoring data can be surprisingly problematic (cf. Inouye et al. 2019).

The most obvious and common approaches to quantifying phenology are to analyze the dates of first, mean, or, more occasionally, last observations (e.g. Roy and Sparks 2000, Forister and Shapiro 2003, Tryjanowski et al. 2005, Fric et al. 2020). Mean dates can be a useful first analysis, but changes in the mean date do not always correspond to changes in first or last events (Inouye et al. 2019). For example, if first and last frost shift progressively earlier and later respectively, a species might become active sooner and stay active later without shifting the mean date of activity. For many types of data sets, observations of first and last events are known to be biased (Strien et al. 2008, Miller-Rushing et al. 2008, Moussus et al. 2010, Bertin 2015, Inouye et al. 2019). Day of first or last observation is very sensitive to population size and sample effort; more data points increase the probability of observing a rare event, so the first observation is earlier, even if the underlying phenological pattern has not actually changed. When a population might be experiencing unmeasured changes in size through time, it becomes impossible to determine whether changes in first/last observation are a consequence of changes in actual phenology, or changes in population size (Miller-Rushing et al. 2008, Pearse et al.

2017).

When the full phenological distribution of dates is available, a large and growing number of methods have been proposed for fitting curves to phenological data. In spite of their potential use, these approaches may be challenging for ecologists because they typically require fairly advanced understanding of statistical computing to implement and interpret, and models with many parameters will require considerable data to parametrize. For example, some ecologists have derived phenological models for insects (Zonneveld et al. 1991, Soulsby and Thomas 2012) and plants (Franco 2018, Steer et al. 2019) from first principles, and used custom-coded methods to fit these to data. Pearse et al (2017) used a Wald distribution to estimate the onset of a phenological events; this method is robust to sample size, and also provides confidence limits of the onset date. Similarly, several studies (Bishop et al. 2013, Dennis et al. 2015, Oke et al. 2019, Stewart et al. 2020) have used custom-coded Gaussian curves to describe the distribution of events through time. Malo (2002) proposes fitting phenology count data with an exponential sine equation; Clark and Thompson (2011) suggest fitting the same type of data with a generalized epsilon-skew-normal curve. For satellite images of vegetation states, the **phenor** package in R (Hufkens et al. 2018) provides a suite of modeling tools. The flexible framework of Generalized Additive Models (GAMs) has been used for a wide range of both univoltine and multivoltine species (e.g. Knudsen et al. 2007, Moussus et al 2009, Hodgson et al. 2011, Strebel et al 2014, Newson et al. 2016, Stemkowski *in press*). At the complex end of available methods, Chapman et al (2015) used Bayesian methods to fit flowering times in Great Britain gardens with a discrete Fourier transform, Park et al (2019) fit data from herbarium specimen with hierarchical Bayesian regression models, and Mizel et al (2019) presented an N-mixture model with more than a dozen parameters that separate spatial, phenological, and population patterns.

Here, we suggest a simple method of modeling phenological distributions using generalized linear mixed models (GLMMs). This method is intermediate in complexity between simple phenological metrics such as first and mean observation dates and custom-coded mechanistic models. In particular, we show how a Gaussian (normal) curve can be fit to phenological data using linear models, including GLMMs. The Gaussian curve is a better starting place than simpler metrics because it captures changes in the duration of phenological activity (unlike mean day of activity) and is not biased by population size or sample effort (unlike first or last observations) (cf. Stewart et al. 2020). Fitting unimodal phenological data with a Gaussian curve is not new (e.g. Clark and Thompson 2011, Bishop et al. 2013, Dennis et al. 2015, Oke et al. 2019, Stewart et al. 2020), but in the past this required custom coding. By fitting Gaussian curves with linear models, it is easy to use out-of-the-box regression methods. We are especially interested in GLMMs because they allow sites or years with extensive data to inform parameter estimates for sites or years with few data. To illustrate this method, we first outline the algebra behind the procedure, then demonstrate its application to an 8-year time series of monitoring from a population of Baltimore checkerspot butterflies (*Euphydryas phaeton*) in Massachusetts.

Gaussian curve as a linear model

The basis of our method is that a Gaussian curve has the form

$$f(x) = ae^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad 1$$

When a is chosen to make Equation 1 integrate to 1, this is the normal or Gaussian distribution. Since everything in Equation 1 is a constant except for x, if we multiply it out and define β_0, β_1, β_2 appropriately in terms of the other constants (see Appendix S1 for the algebra), we can rewrite

the Gaussian curve as

$$f(x) = e^{\beta_0 + \beta_1 x + \beta_2 x^2} \qquad 2$$

Here we can see that the terms in the exponent are a quadratic equation. This means that if we take the natural log of both sides, we are left with a familiar linear model with both a linear and a quadratic term:

$$\ln(y) = \beta_0 + \beta_1 x + \beta_2 x^2 \qquad 3$$

Despite our special use for it, Equation 3 is an ordinary linear model of a quadratic equation, and can be fit with standard tools for linear models. In the context of phenology, the most straightforward analysis would use empirical estimates of abundance or activity (e.g. transect counts of butterflies or flowers) for dependent variable *y*, which is distributed following a Gaussian curve in relation to some measure of time (e.g. day of year) for the independent variable *x* (Fig. 1A-H).

Fitting a linear model of *ln(y)* vs. *x* provides estimates and confidence intervals for β_0, β_1, and β_2. By reversing the algebra between Equations 1 and 2, we can recover the parameters of the Gaussian curve (mean μ, variance σ^2), as well as metrics determined by the gaussian curve (e.g. area under the curve) that may be useful in interpreting phenological data (Appendix S1: Fig. S1). First, μ, the estimated **day of peak activity** and **mean day of activity** (these are the same since the Gaussian curve is symmetrical) can be calculated from the slopes of the linear and quadratic terms:

$$\mu = -\frac{\beta_1}{2\beta_2} \qquad 4$$

The standard deviation of the Gaussian curve, σ, is a function of the slope of the quadratic term:

$$\sigma = \sqrt{-\frac{1}{2\beta_2}} \qquad 5$$

In the case study below, we refer to 2*1.285*σ as the **flight period**. This corresponds to the range of dates between the 0.1 and the 0.9 quantiles, and so this measures the duration of time when the middle 80% of observations are estimated to occur (Jonzén et al. 2006, Michielini et al. *in review*).

The area under the gaussian curve, N, is a **population abundance index:**

$$N = \left(\sqrt{-\frac{2\pi}{2\beta_2}}\right)\exp\left(\beta_0 - \frac{{\beta_1}^2}{4\beta_2}\right) \qquad 6$$

We use the term "abundance index" because, for constant sampling effort and detection probability, *N* will be proportional to the number of active individuals. However, it is actually a measure of estimated activity-days (Dennis et al. 2015, Wepprich et al. 2019).

We note that here we have focused on a few phenological metrics, including the days of 0.1, 0.5 (e.g. day of peak) and 0.9 quantile. However, because our proposed approach fits a Guassian curve, from estimated μ and σ it is trivial to calculate any metrics of a Gaussian curve, including (a) any arbitrary quantiles (e.g. the 0.05 and 0.95 quantiles used in Stemkowski et al *in Press*), (b) the height of the curve (e.g. maximum number of flowers, Miller-Rushing and Inouye (2009)), or (c) the "observable flight season" (days when the curve exceeds 1, a metric reflecting the period of likely human detection that parallels first and last observation dates).

Standard errors of derived parameters such as μ, σ^2, or N can be estimated using the delta method (Williams *et al.* 2002), or by parametric bootstrapping (Dennis 1996). Code for these analyses is given in Appendices S2 (tutorial as html), S3 (analysis as html) and S4 (data and Rmarkdown sources for Appendices S2 and S3). All code was written and run in R version 4.0.0

(R Core Team, 2020).

Case study

Data set

From 2012-2019, we conducted a capture-recapture study of Baltimore checkerspot (*Euphydryas phaeton*) butterflies at a natural area (Williams Conservation Land) in the town of Harvard MA, USA (Brown and Crone 2016, Brown et al. 2017, Crone 2018). Baltimore checkerspot is a univoltine species, with one clear flight period of adults per year. Surveys were conducted by visiting the site 2-3 times a week from mid-June until the population was clearly finished for the year; the onset of checkerspot flight at this site is usually in late June or early July. To illustrate the use of a Gaussian curve to estimate phenological metrics, we converted capture-recapture data to counts of individual animals handled on each visit to the site. This monitoring protocol creates a data structure that is similar to traditional "Pollard walk" style monitoring (Pollard 1977, Pollard and Yates 1993, Wepprich et al. 2019) but differs from Pollard walks in that the site was searched freely, rather than by walking a fixed route. For comparison with the Gaussian GLMM analyses below, we estimated population size each year using standard open population capture-recapture models (see supplemental methods).

Estimation of phenology metrics

Methods: We fit Gaussian curves to these data using Poisson family, log link GLMMs, with the number of butterflies seen on each day as the dependent variable, day of year and day of year squared as independent variables, and random effects of the intercept, day of year, and day of year squared terms with respect to year. Note that the log-linear relationship (eq. 3) is captured

by the log-link model. We also include an observation-level random effect (Elston et al. 2001) to account for overdispersion (ΔAIC = 263.7 compared to model fit without this term). After fitting the linear model, we used equations 4-7 to calculate the estimated mean day of activity, standard deviation of flight period, population abundance, and peak abundance for each year, and use mean day of activity and standard deviation of activity to calculate onset (day of 0.1 quantile) and end (day of 0.9 quantile) of activity. For heuristic purposes, we calculated confidence intervals for these metrics using both the delta method and parametric bootstrapping.

Results: For this univoltine butterfly population, the Gaussian curve provides a visually satisfying fit, with the expected property that years with few surveys (2018 and 2019, Fig. 1G-H) are informed by years with more surveys (Fig. 1 A-F). Estimates of the mean date of activity and the flight period were reasonably precise, and adequate to detect among-year variation (Figs. 1I-J). Estimates of the abundance index were noticeably imprecise (Fig. 1K). However, in spite of large uncertainty for estimates of abundance index, there is a strong correspondence between abundance index and population estimates from the capture-recapture study ($R^2 = 0.94$) (Fig. 2A). These data also show the bias of first and last dates of observations in relation to population size (Fig. 2B-C); compared to 0.1 and 0.9 quantiles estimated from annual Gaussian curves, years with smaller populations had later first observations and earlier last observations.

Evaluating an environmental driver

Methods: Given that phenological events change through time, a common next question is to ask whether these changes are associated with changes in environmental conditions (see, e.g., Roy and Sparks 2000, Forister and Shaprio 2003, Marra et al. 2005, Jonzén et al. 2006, Miller-Rushing et al. 2008, van Buskirk et al. 2009, Hodgson et al. 2011, Gordo et al. 2013,

Bertin 2015, Cayton et al. 2015, Barton and Sandercock 2017, Heberling et al. 2019, Oke et al. 2019, Park et a. 2019, Fric et al. 2020, Horton et al. 2020, Stewart et al. 2020, Stemkowski et al. *in press*). It is possible in principle to simultaneously fit drivers of phenological change and parameters themselves (e.g. Mizel et al. 2019). However, the algebra of converting a Gaussian curve to a linear model does not enable easy inclusion of covariates of the ecologically meaningful derived metrics such as onset of activity or peak dates (unpubl. calculations). One accessible alternative to custom-coding complex models is a two-step process of first estimating parameters with GLMMs, then using these derived parameters in subsequent models. To account for uncertainty in derived parameters, it is straightforward to use parameteric bootstrapping.

To illustrate this approach, we compared yearly estimates of the day of mean activity, flight period, and population abundance to temperature. To represent temperature, we used growing degree days (GDD), a measure of thermal environment that has been found to predict plant and insect phenology (Hodgson et al. 2011, Cayton et al. 2015) (see Appendix S1). For each phenological metric, we fit a simple linear regression with GDD as the predictor. We also calculated 95% confidence intervals for the slope using parametric bootstrapping, and the proportion of p values that were less than 0.05 among these bootstrapped model fits.

Results: Mean day of activity was consistently earlier in warmer years (bootstrapped 95% CI of slope of μ vs. GDD: [-0.088, -0.21]), but it was rarely statistically significant ($p<0.05$ only 11% of the time) (Fig. 3B). Temperature was not associated with differences in flight period (bootstrapped 95% CI of σ vs. GDD: [-0.026, 0.50], $p<0.05$ only 2% of the time) or abundance index (bootstrapped CI of [-18.54, 2.60], $p<0.05$ less than 0.01% of the time) (Fig. 3C-D). Because the mean day of activity advanced with GDD but the flight period was not associated with GDD, it is not surprising that the onset and end of activity also advanced with GDD in this

population (bootstrapped 95% CI slopes of [-0.95, -0.02] and [-0.08, -0.01] for onset (0.1 quantile) and end (0.9 quantile) respectively, $p<0.05$ in 35% and 2% of bootstrap samples, respectively).

Discussion

In this paper, we show that a Gaussian curve can be fit to phenological data using familiar methods for linear models. We hope this approach fills a gap in the complexity gradient of available tools for analyzing phenological patterns. Currently, objectively problematic metrics for phenological patterns such as first observations are still used by at least some ecologists. Of course, sometimes data limitations constrain analysis to only use first or last metrics, especially when comparing with historic data sets (e.g. Heberling et al. 2019). However, in many cases ecologists have much more complete data, and should not be limited to using problematic phenological metrics. Fitting Gaussian curves with GLMMs may be a reasonable first step for many ecologists interested in describing phenology.

Although it is possible to fit Gaussian curves in the familiar framework of GLMMS, this method uses GLMMs in an unfamiliar way. Therefore, some features of this differ from typical modeling conventions. For example, in this context, the exact parameterization of the GLMM is important for its ecological meaning. If the goal is to fit Gaussian curves with different ecological parameters to each year of data, it may not be appropriate to include or exclude linear and quadratic phenology terms from fixed and random effects based on significance testing. Another feature of mixed models that may be unfamiliar to some ecologists is shrinkage. Shrinkage is a general property of mixed and hierarchical models, in that values for each year (or sampling unit) are biased towards the mean, especially when sample size is low (Gelman et al.

2003; ch 2 & 5). This feature is a desirable property, in the sense that it balances signal and noise from years with fewer observations (see Fig. 1G). In our analysis of environmental covariates, however, shrinkage could reduce our ability to detect covariate effects (Fig. 3B-D), because derived metrics are pulled toward average values before covariates are taken into account. It would be statistically more powerful – but computationally more difficult – to simultaneously estimate both covariate effects and phenological models in one custom-coded model (see, e.g., Mizel et al. 2019).

The combination of a simple mathematical form (3 parameters) and shrinkage allows Gaussian GLMMs to estimate phenology and abundance in years with relatively few observations (see, e.g., Fig 1H). This ability may be important in studies of at-risk species or declining populations. Recent studies of butterfly (Wepprich et al. 2019) and bee (Stemkovski et al. *in press*) populations have been restricted to relatively abundant species by the needs of sufficient data to fit abundances within years using very flexible functions. Similarly, in their analyses with more complex phenology models, Zonneveld (1991) had between 14 and 30 records per year, Clark and Thompson (2011) had between 24 and 61 records per year, Knudsen et al (2007) had 39 records in their demonstration year, Moussus et al. (2008) had between 715 and 3544 records per year, and Strebel et al. (2015) excluded years with fewer than 20 observations. In many cases, hierarchical Bayesian models also incorporate shrinkage (Gelman et al. 2005). However, recent of implementations of hierarchical models in phenological research (e.g. Chapman et al 2015, Park et al 2019) still have a relatively large number of parameters, so may not be as robust to small sample sizes as the simple 3-parameter Gaussian curve.

In the Baltimore checkerspot case study, confidence intervals were modest for

phenological metrics, but large for abundance. This result may be a feature of the approach in general: because the formula for abundance index, *N* (eq 6), involves exponentiating coefficients on top of the multiplication and division involved in the other phenological metrics, any uncertainty in the coefficients will lead to dramatically increased uncertainty in N. Abundance has also been difficult to estimate precisely from transect data using other models (Gross et al. 2007). In spite of relatively large error, the tight correlation between N, the abundance index from this model, and capture-recapture estimates of population size calculated separately from the same data (Fig. 2) suggests that abundance estimates are unbiased. The correlation also compares favorably with other methods of fitting transect data; Haddad et al. (2008) found no correlation between mark-recapture estimates of population size and population size estimated using the Zonneveld model. However, the fact that we find a 1:1 match of *N* (an index that reflects longevity as well as abundance) and capture-recapture estimates is likely coincidental. By chance, our capture probability during surveys was ≈ 0.15 (unpubl.), and the apparent survival of Baltimore checkerspot butterflies at our site is 0.844/day (Brown and Crone 2016); these values mean that in our example, the capture probability - by chance - exactly cancelled out the fact that *N* is actually in units of "butterfly days".

Gaussian curves are only well-suited to represent data that is unimodal and approximately symmetric. For many phenological events, we suspect the assumption of symmetry is a reasonable approximation (see Fig 1, and Stewart et al. 2020), although this is of course a hypothesis that could be explored depending on the goals of an analysis. Multimodal distributions may be more problematic. For multivoltine insects, Generalized Additive Models (GAMs) (e.g. Knudsen et al. 2007, Moussus et al 2009, Hodgson et al. 2011, Newson et al. 2016, Stemkovski *in press*) have been used to capture changes in phenology over time. In this kind of

application, GAMs and linearized-Gaussian GLMMs share the feature of fitting a statistical model to the distribution of events in time. Although they are not described by a parametric equation, features like the onset (0.1 quantile) or end (0.9 quantile) can be extracted from GAMs numerically (cf. Stemkovski *et al. in press*). Another approach to evaluating phenological events without assuming a particular distribution is quantile regression (Cade and Noon 2003, Koenker 2019), which has been used in several studies of bird migration (e.g. Gordo et al 2013, Barton and Sandercock 2018), and occasionally for Lepidoptera (Gimesi et al. 2012, Michielini et al. *in review*). Quantile regression shares the property of drawing on well-established and well-validated statistical approaches, rather than developing new ones.

In spite of some limitations in analytical methods to date, changes in phenology are apparent in many ecological systems. Because they are one of the most conspicuous signs of climate change, there is naturally interest in the causes and consequences of such shifts. We expect that the widespread interest in phenology will continue to lead to a growing number of new methods for interpreting patterns in phenological data. At the same time, not every new method is guaranteed to work for all (or even most) data, and custom-coding for every question can be time consuming, prone to errors, and intimidating to many ecologists. We encourage ecologists to be aware of well-established existing methods – like the linearized Gaussian model we present here – that may be useful for evaluating phenological patterns in a robust way.

Declarations: This work was conducted with support from NSF DEB 19-20834 and the DOD SERDP program (RC-2700). We thank many Crone lab members who participated in Baltimore checkerspot data collection, especially Lee Brown, and Brian Inouye for helpful suggestions on the manuscript. Data and code will be available on Dryad.

References:

Barton, G. G., and B. K. Sandercock. 2018. "Long-term changes in the seasonal timing of landbird migration on the pacific flyway." *The Condor: Ornithological Applications* 120(1): 30-46.

Bates, Douglas, Martin Maechler, Ben Bolker, Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1): 1-48.

Bertin, Robert I. 2015. "Climate Change and Flowering Phenology in Worcester County, Massachusetts." *International Journal of Plant Sciences* 176 (2): 107–19.

Bishop, Tom R., Marc S. Botham, Richard Fox, Simon R. Leather, Daniel S. Chapman, and Tom H. Oliver. 2013. "The Utility of Distribution Data in Predicting Phenology." *Methods in Ecology and Evolution* 4(11): 1024–32.

Bolker, Ben and others. 2020. "GLMM FAQ". https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html.

Brown, Leone M., and Elizabeth E. Crone. 2016. "Minimum area requirements for an at-risk butterfly based on movement and demography: movement, demography, and minimum area." *Conservation Biology* 30(1): 103–112.

Brown, Leone M., Rebecca K. Fuda, Nicolas Schtickzelle, Haley Coffman, Audrey Jost, Alice Kazberouk, Eliot Kemper, Emma Sass, and Elizabeth E. Crone. 2017. "Using Animal Movement Behavior to Categorize Land Cover and Predict Consequences for Connectivity and Patch Residence Times." *Landscape Ecology* 32 (8): 1657–70.

Cade, Brian S., and Barry R. Noon. 2003. "A gentle introduction to quantile regression for ecologists." *Frontiers in Ecology and the Environment* 1(8): 412–420.

Cayton, Heather L., Nick M. Haddad, Kevin Gross, Sarah E. Diamond, and Leslie Ries. 2015. "Do growing degree days predict phenology across butterfly species?" *Ecology* 96(6): 1473-

1479.

Chapman, Daniel S., Sandra Bell, Stephan Helfer, and David B. Roy. 2015. “Unbiased inference of plant flowering phenology from biological recording data.” *Biological Journal of the Linnean Society* 115(3):543–554.

Chmura, Helen E., Heather M. Kharouba, Jaime Ashander, Sean M. Ehlman, Emily B. Rivest, and Louie H. Yang. 2019. “The mechanisms of phenology: the patterns and processes of phenological shifts.” *Ecological Monographs* 89(1):e01337.

Clark, Robert M., and Roy Thompson. 2011. “Estimation and Comparison of Flowering Curves.” *Plant Ecology & Diversity* 4 (2-3): 189–200.

Cleland, Elsa E., Isabelle Chuine, Annette Menzel, Harold A. Mooney, and Mark D. Schwartz. 2007. “Shifting Plant Phenology in Response to Global Change.” *Trends in Ecology & Evolution* 22 (7): 357–65.

Crone, E. 2018. “The rise and fall of a checkerspot population.” *News Lepid. Soc* 60:40–41.

Dennis, Brian. 1996. “Discussion: Should Ecologists Become Bayesians?” Ecological Applications 6(4):1095–1103.

Dennis, Emily B., Stephen N. Freeman, Tom Brereton, and David B. Roy. 2013. “Indexing butterfly abundance whilst accounting for missing counts and variability in seasonal pattern.” *Methods in Ecology and Evolution* 4(7): 637–645.

Dennis, Emily B., Byron J. T. Morgan, Stephen N. Freeman, David B. Roy, and Tom Brereton. 2015. “Dynamic models for longitudinal butterfly data.” *Journal of Agricultural, Biological, and Environmental Statistics* 21(1): 1–21.

Elston, D. A., R. Moss, T. Boulinier, C. Arrowsmith, and X. Lambin. 2001. “Analysis of aggregation, a worked example: numbers of ticks on red grouse chicks.” *Parasitology*

122(5): 563–569.

Forister, Mathew, and Arthur Shapiro. 2003. “Climatic Trends and Advancing Spring Flight of Butterflies in Lowland California.” *Global Change Biology* 9(7): 1130-1135.

Fric, Faltýnek, Michal Rindoš, and Martin Konvička. 2020. “Phenology Responses of Temperate Butterflies to Latitude Depend on Ecological Traits.” *Ecology Letters* 23 (1): 172–80.

Franco, Miguel. 2018. “The time distribution of biological phenomena – illustrated with the London marathon.” PeerJ Preprints e27175v1

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2003. *Bayesian data analysis*. CRC press.

Gimesi, László, Réka Homoródi, Anikó Hirka, Károly Szenteleki, and Levente Hufnagel. 2012. “The effect of climate change on light trap capturing data of Lepidoptera.” International *Proceedings of Chemical, Biological and Environmental Engineering* 28: 255–259.

Gordo, Oscar, and Juan José Sanz. 2009. “Long-Term Temporal Changes of Plant Phenology in the Western Mediterranean.” *Global Change Biology* 15(8): 1930–48.

Gordo, Oscar, Piotr Tryjanowski, Jakub Z. Kosicki, and Miroslav Fulín. 2013. “Complex phenological changes and their consequences in the breeding success of a migratory bird, the white stork *Ciconia ciconia*.” *The Journal of Animal Ecology* 82(5): 1072–1086.

Gross, Kevin, Eric J. Kalendra, Brian R. Hudgens, and Nick M. Haddad. 2007. “Robustness and uncertainty in estimates of butterfly abundance from transect counts.” *Population Ecology* 49(3): 191–200.

Haddad, Nick M., Brian Hudgens, Chris Damiani, Kevin Gross, Daniel Kuefler, and Ken Pollock. 2008. “Determining Optimal Population Monitoring for Rare Butterflies: Monitoring Rare Butterflies.” *Conservation Biology* 22(4): 929–40.

Heberling, J. Mason, Caitlin McDonough MacKenzie, Jason D. Fridley, Susan Kalisz, and

Richard B. Primack. 2019. “Phenological mismatch with trees reduces wildflower carbon budgets.” *Ecology Letters* 22(4): 616–623.

Hodgson, J. A., C. D. Thomas, T. H. Oliver, B. J. Anderson, T. M. Brereton, and E. E. Crone. 2011. “Predicting insect phenology across space and time.” *Global Change Biology* 17(3): 1289–1300.

Horton, Kyle G., Frank A. La Sorte, Daniel Sheldon, Tsung-Yu Lin, Kevin Winner, Garrett Bernstein, Subhransu Maji, Wesley M. Hochachka, and Andrew Farnsworth. 2020. “Phenology of Nocturnal Avian Migration Has Shifted at the Continental Scale.” *Nature Climate Change* 10(1): 63–68.

Hufkens, Koen, David Basler, Tom Milliman, Eli K. Melaas, and Andrew D. Richardson. 2018. “An Integrated Phenology Modelling Framework in R.” *Methods in Ecology and Evolution* 9(5): 1276–85.

Inouye, Brian D., Johan Ehrlén, and Nora Underwood. 2019. “Phenology as a process rather than an event: from individual reaction norms to community metrics.” *Ecological Monographs* 89(2): e01352.

Jonzén, Niclas, Andreas Lindén, Torbjørn Ergon, Endre Knudsen, Jon Olav Vik, Diego Rubolini, Dario Piacentini, et al. 2006. “Rapid advance of spring arrival dates in long-distance migratory birds.” *Science* 312(5782): 1959–1961.

Knudsen, E., A. Lindén, T. Ergon, N. Jonzén, J. O. Vik, J. Knape, J. E. Røer, and N. C. Stenseth. 2007. “Characterizing bird migration phenology using data from standardized monitoring at bird observatories.” *Climate Research* 35: 59–77.

Koenker, Roger. 2019. “Quantreg: Quantile Regression.” https://CRAN.R-project.org/package=quantreg.

Malo, J. E. 2002. “Modelling Unimodal Flowering Phenology with Exponential Sine Equations.” *Functional Ecology* 16(3): 413–18.

Marra, Peter P., Charles M. Francis, Robert S. Mulvihill, and Frank R. Moore. 2005. “The influence of climate on the timing and rate of spring bird migration.” *Oecologia* 142(2): 307–315.

Michielini, J., E. Dopman, and E. E. Crone. In review at *Ecology Letters*. “Changes in flight period predict trends in Massachusetts butterflies.”

Miller-Rushing, Abraham J., David W. Inouye, and Richard B. Primack. 2008. “How well do first flowering dates measure plant responses to climate change? The effects of population size and sampling frequency.” *Journal of Ecology* 96(6): 1289-1296.

Mizel, Jeremy D., Joshua H. Schmidt, Laura M. Phillips, and Carol L. Mcintyre. 2019. “A binomial N□mixture model for estimating arrival and departure timing.” *Methods in Ecology and Evolution* 10(7): 1062–71.

Moussus, J. P., F. Jiguet, J. Clavel, and R. Julliard. 2009. “A Method to Estimate Phenological Variation Using Data from Large□scale Abundance Monitoring Programmes.” *Bird Study: The Journal of the British Trust for Ornithology* 56(2): 198-212.

Moussus, Jean-Pierre, Romain Julliard, and Frédéric Jiguet. 2010. “Featuring 10 Phenological Estimators Using Simulated Data: Featuring the Behaviour of Phenological Estimators.” *Methods in Ecology and Evolution* 1(2): 140–50.

Newson, Stuart E., Nick J. Moran, Andy J. Musgrove, James W. Pearce-Higgins, Simon Gillings, Philip W. Atkinson, Ryan Miller, Mark J. Grantham, and Stephen R. Baillie. 2016. “Long-Term Changes in the Migration Phenology of UK Breeding Birds Detected by Large-Scale Citizen Science Recording Schemes.” *The Ibis* 158(3): 481–95.

Oke, Krista B., Curry J. Cunningham, Thomas P. Quinn, and Andrew P. Hendry. 2019. "Independent lineages in a common environment: the roles of determinism and contingency in shaping the migration timing of even-versus odd-year pink salmon over broad spatial and temporal scales." *Ecology Letters* 22(10): 1547–1556.

Park, Daniel S., Ian Breckheimer, Alex C. Williams, Edith Law, Aaron M. Ellison, and Charles C. Davis. 2019. "Herbarium Specimens Reveal Substantial and Unexpected Variation in Phenological Sensitivity across the Eastern United States." *Philosophical Transactions of the Royal Society of London* 374(1763): 20170394.

Pearse, William D., Charles C. Davis, David W. Inouye, Richard B. Primack, and T. Jonathan Davies. 2017. "A statistical estimator for determining the limits of contemporary and historic phenology." *Nature Ecology & Evolution* 1(12):1876-1882.

Pollard, E. 1977. "A method for assessing changes in the abundance of butterflies." *Biological Conservation* 12(2):115–134.

Pollard, E., and T. J. Yates. 1993. *Monitoring Butterflies for Ecology and Conservation*. Chapman & Hall, London.

R Core Team. 2020. "R: A language and environment for statistical computing". R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Roy, D. B., and T. H. Sparks. 2000. "Phenology of British butterflies and climate change." *Global Change Biology* 6(4):407–416.

Soulsby, R.L. and Thomas, J.A. 2012. "Insect population curves: modelling and application to butterfly transect data." *Methods in Ecology and Evolution* 3(5): 832-841.

Steer, Nicola C., Paul M. Ramsay, and Miguel Franco. 2019. "Nlstimedist: an r package for the biologically meaningful quantification of unimodal phenology distributions." *Methods in*

Ecology and Evolution 10(11): 1934–1940.

Strebel, Nicolas, Marc Kéry, Michael Schaub, and Hans Schmid. 2014. "Studying Phenology by Flexible Modelling of Seasonal Detectability Peaks." *Methods in Ecology and Evolution*, 5(5): 483–90.

Stemkowski, *in press*.

Stewart, James E., Javier Gutiérrez Illán, Shane A. Richards, David Gutiérrez, and Robert J. Wilson. 2020. "Linking inter□annual variation in environment, phenology, and abundance for a montane butterfly community." *Ecology* 101(1): e02906.

Strien, Arco J. van, Willem F. Plantenga, Leo L. Soldaat, Chris A. M. van Swaay, and Michiel F. Wallisdevries. 2008. "Bias in phenology assessments based on first appearance data of butterflies." *Oecologia* 156(1): 227–235.

Tryjanowski, Piotr & Kuźniak, S. & Sparks, Tim. 2005. "What Affects the Magnitude of Change in First Arrival Dates of Migrant Birds?" *Journal of Ornithology* 146(3): 200–205.

van Buskirk, Josh, Robert S. Mulvihill, and Robert C. Leberman. 2009. "Variable Shifts in Spring and Autumn Migration Phenology in North American Songbirds Associated with Climate Change." *Global Change Biology* 15(3): 760–71.

Wepprich, Tyson, Jeffrey R. Adrion, Leslie Ries, Jerome Wiedmann, and Nick M. Haddad. 2019. "Butterfly abundance declines over 20 years of systematic monitoring in Ohio, USA." *PloS One* 14(7): e0216270

Williams, Byron K., James D. Nichols, and Michael J. Conroy. 2002. *Analysis and management of animal populations*. Academic Press.

Zonneveld, Cor. 1991. "Estimating death rates from transect counts." *Ecological Entomology* 16(1): 115–121.

Figure 1: Gaussian model fitted to Baltimore checkerspot butterfly data. (A-H) Points show raw data, blue lines show best-fitting Gaussian curve, dashed gray lines show +/- 1 standard error. For comparability, day of month in this and other figures is based on a 365 day year (excludes leap days). Note the different scales on the y-axes. (I-K) fitted metrics: flight period (I), mean day of activity (J), and abundance index (C). Bars show +/- 1 standard error, calculated using the delta method (black) or parametric bootstrapping (blue). In all panels, parametric bootstrap standard errors are 0.16 and 0.84 quantiles of bootstrap samples.

Figure 2: Evaluation of Gaussian GLMMs. (A) Comparison of estimated abundance index to mark-recapture estimates. Black lines show +/- 1 standard error of each metric, dashed line shows correlational ellipse. (B-C) Comparison of quantile-based estimates of onset (B) and end (C) of activity to day of first (B) and last (C) observation for each year, with point size scaled by population size (square-root scale). Black lines show a 1:1 line going through the mean of observed values. As expected, years with smaller population sizes show less extreme observations (i.e., points above (B) and below (C) the 1:1 line).

Figure 3: (A) Fitted Gaussian curves of activity for each year, colored by growing degree days. Post-hoc analysis of the effects of yearly cumulative growing degree days (GDD) calculated on July 1 with (B) mean day of activity, (C) Flight period (number of days between the 0.1 and 0.9 quantiles), and (D) abundance index. Bars show +/- 1 standard error of the annual estimates, calculated through parametric bootstrapping. Black lines show the relationship of the means, and dashed lines show the 0.16-0.84 quantiles of bootstrapped analyses (analogous to +/- 1 standard error of the regressions).

Figure 1.

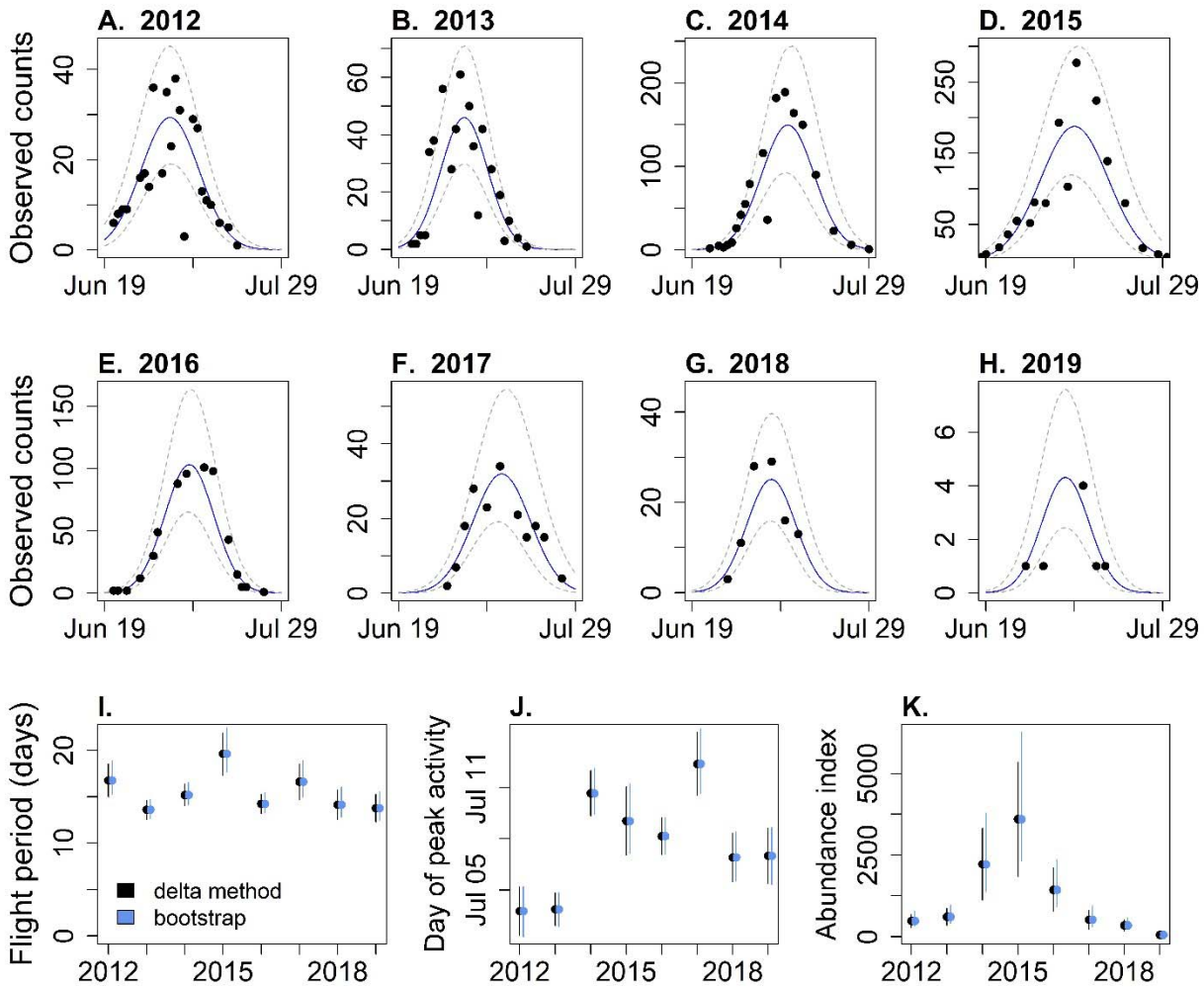

A. 2012
B. 2013
C. 2014
D. 2015
E. 2016
F. 2017
G. 2018
H. 2019
Observed counts
Jun 19
Jul 29
I.
J.
K.
Flight period (days)
Day of peak activity
Abundance index
Jul 05
Jul 11
delta method
bootstrap
2012
2015
2018

Figure 2

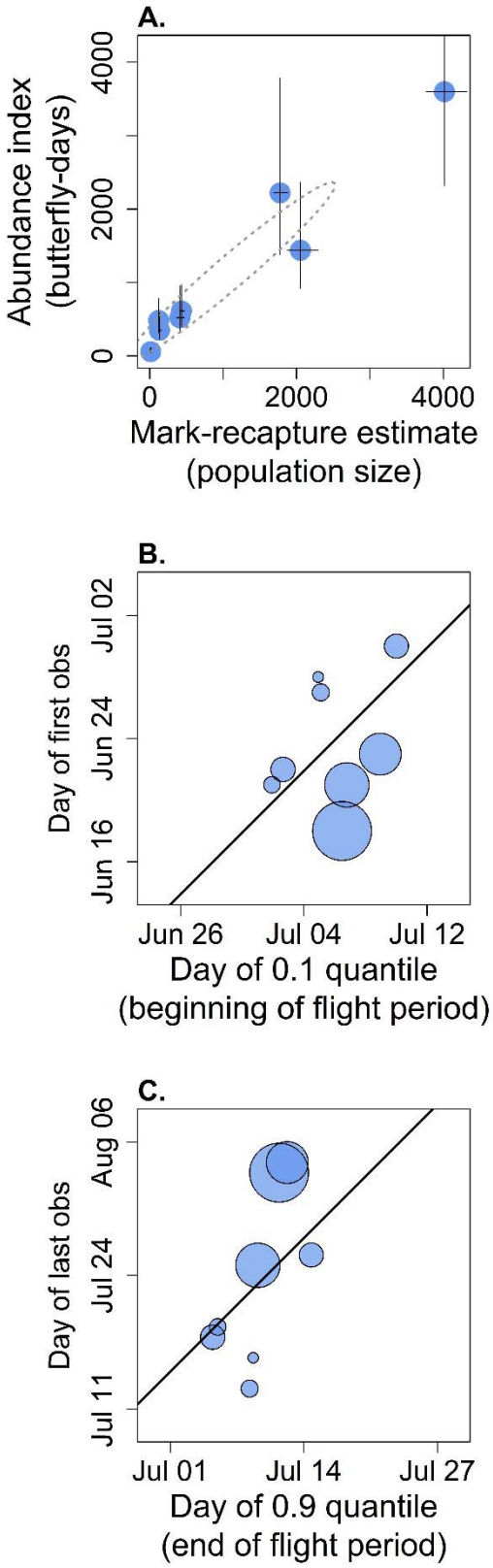
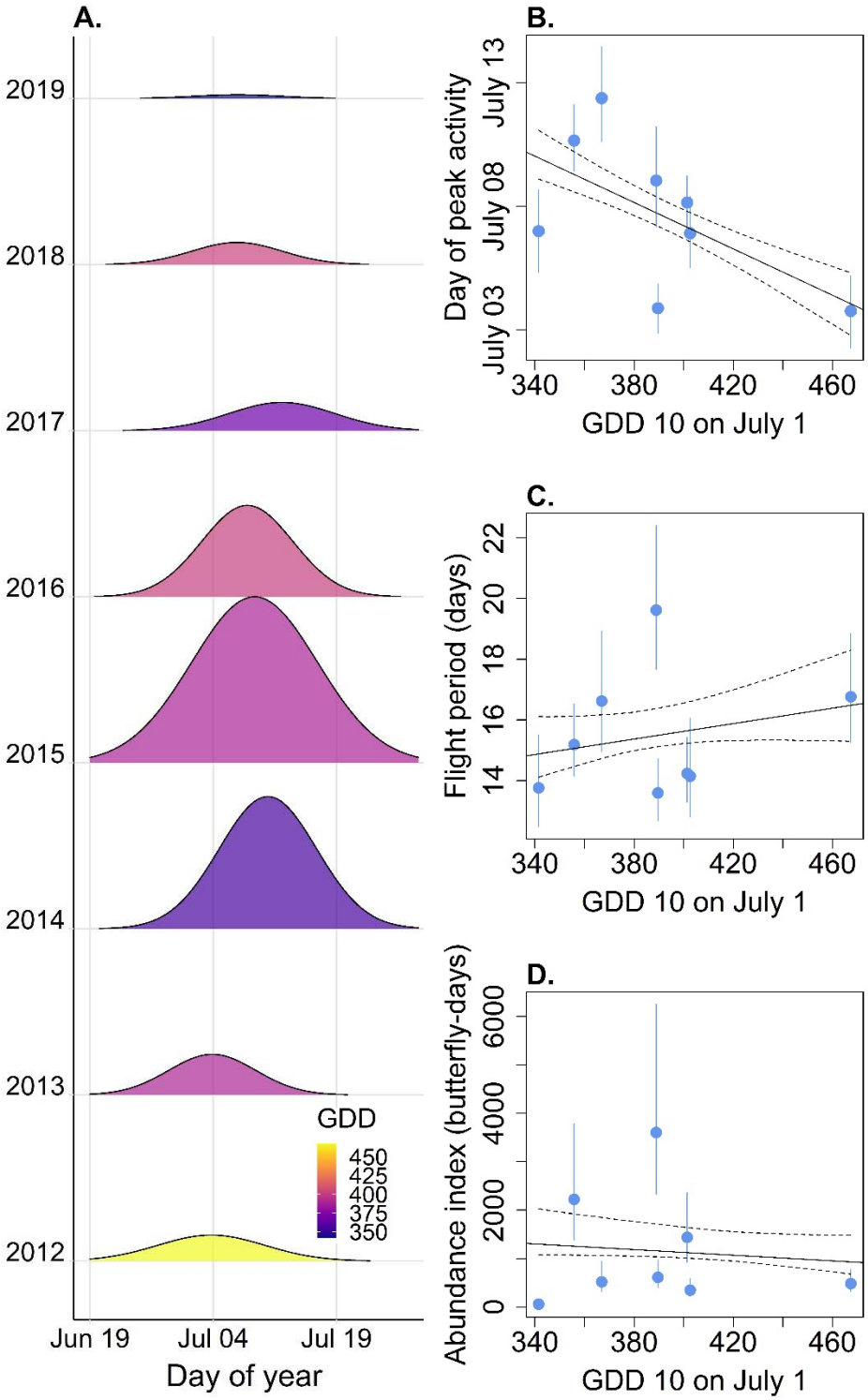

A.
Abundance index
(butterfly-days)
0
2000
4000
0
2000
4000
Mark-recapture estimate
(population size)
B.
Day of first obs
Jun 16
Jun 24
Jul 02
Jun 26
Jul 04
Jul 12
Day of 0.1 quantile
(beginning of flight period)
C.
Day of last obs
Jul 11
Jul 24
Aug 06
Jul 01
Jul 14
Jul 27
Day of 0.9 quantile
(end of flight period)

Figure 3:


A.
2019
2018
2017
2016
2015
2014
2013
2012
GDD
450
425
400
375
350
Jun 19
Jul 04
Jul 19
Day of year
B.
Day of peak activity
July 03
July 08
July 13
340
380
420
460
GDD 10 on July 1
C.
Flight period (days)
14
16
18
20
22
340
380
420
460
GDD 10 on July 1
D.
Abundance index (butterfly-days)
0
2000
4000
6000
340
380
420
460
GDD 10 on July 1