

# Kent Academic Repository

## Full text document (pdf)

### Citation for published version

Dennis, Emily B. and Morgan, Byron J. T. and Freeman, Stephen N. and Roy, David B. and Brereton, Tom (2015) Dynamic models for longitudinal butterfly data. *Journal of Agricultural, Biological, and Environmental Statistics*, 21 (1). pp. 1-21. ISSN 1085-7117.

### DOI

<https://doi.org/10.1007/s13253-015-0216-3>

### Link to record in KAR

<http://kar.kent.ac.uk/46264/>

### Document Version

Publisher pdf

#### Copyright & reuse

Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

#### Versions of research

The version in the Kent Academic Repository may differ from the final published version.

Users are advised to check <http://kar.kent.ac.uk> for the status of the paper. **Users should always cite the published version of record.**

#### Enquiries

For any further enquiries regarding the licence status of this document, please contact:

[researchsupport@kent.ac.uk](mailto:researchsupport@kent.ac.uk)

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at <http://kar.kent.ac.uk/contact.html>

# Dynamic Models for Longitudinal Butterfly Data

Emily B. DENNIS, Byron J. T. MORGAN, Stephen N. FREEMAN, David B. ROY, and Tom BRERETON

We present models which provide succinct descriptions of longitudinal seasonal insect count data. This approach produces, for the first time, estimates of the key parameters of brood productivities. It may be applied to univoltine and bivoltine species. For the latter, the productivities of each brood are estimated separately, which results in new indices indicating the contributions from different generations. The models are based on discrete distributions, with expectations that reflect the underlying nature of seasonal data. Productivities are included in a deterministic, auto-regressive manner, making the data from each brood a function of those in the previous brood. A concentrated likelihood results in appreciable efficiency gains. Both phenomenological and mechanistic models are used, including weather and site-specific covariates. Illustrations are provided using data from the UK Butterfly Monitoring Scheme, however the approach is perfectly general. Consistent associations are found when estimates of productivity are regressed on northing and temperature. For instance, for univoltine species productivity is usually lower following milder winters, and mean emergence times of adults for all species have become earlier over time, due to climate change. The predictions of fitted dynamic models have the potential to improve the understanding of fundamental demographic processes. This is important for insects such as UK butterflies, many species of which are in decline. Supplementary materials for this article are available online.

**Key Words:** Abundance indices; Auto-regression; Concentrated likelihood; Generalised additive models; Phenology; Stopover models.

## 1. INTRODUCTION

Climate change is predicted to become an increasingly important cause of biodiversity decline (Thomas et al. 2004; Pereira et al. 2010). Species' responses to climate are often complex and present challenges for modelling and prediction. We illustrate the models of this paper with reference to butterflies, the most comprehensively monitored insects. Their

---

Emily B. Dennis (✉), Byron J. T. Morgan School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, UK. (E-mail: [E.B.Dennis@kent.ac.uk](mailto:E.B.Dennis@kent.ac.uk)). Emily B. Dennis, Tom Brereton Butterfly Conservation, Manor Yard, East Lulworth, Wareham, UK. Stephen N. Freeman, David B. Roy Centre for Ecology & Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford, UK.

© 2015 International Biometric Society  
*Journal of Agricultural, Biological, and Environmental Statistics*, Volume 21, Number 1, Pages 1–21  
DOI: [10.1007/s13253-015-0216-3](https://doi.org/10.1007/s13253-015-0216-3)

population status is increasingly recognised as an indicator for changes in biodiversity as they respond sensitively and rapidly to changes in habitat and climate (Thomas 2005).

Previous studies of UK butterflies imply positive associations of populations with warm summer weather, but predicted relationships with winter weather are variable (Roy et al. 2001; Dennis and Sparks 2007; Isaac et al. 2011). Evidence for shifts in phenology (Roy and Sparks 2000) and increases in voltinism (Altermatt 2010) have also been presented.

Extensive sources of citizen-science count data for butterflies are available both in the UK and around the world, and there is much interest in developing robust modelling approaches to assist the monitoring and understanding of species' responses to change. Butterflies have multi-stage life cycles, and counts fluctuate within each year in response to their emergence as adults, which is generally the only life stage with widespread data. Soulsby and Thomas (2012) developed a mathematical model for this variation, but only allowed for discrete, non-overlapping generations. Other models have been proposed to describe the within-year variation, both non-parametrically using generalised additive models (GAMs, Rothery and Roy 2001; Dennis et al. 2013) and via stochastic mixture models (Matechou et al. 2014; Dennis et al. 2014). Counts adjusted for seasonal fluctuations can then be used to produce longer-term trends, but existing methods do not impose any relationship between counts from one year to the next, which is the topic of this paper.

Causes of variation in both abundance and seasonal pattern from one year to the next are multi-faceted, relating to numbers during the previous year, as well as other factors driving the unobserved stages of the life-cycle, such as weather. We describe a novel dynamic framework which models count data across multiple sites from consecutive years, with abundance in any given year driven by that in the previous year. We adapt the approach for bivoltine species, with the first brood in a year feeding into the second. The models can be fitted efficiently using concentrated likelihoods. Performance is illustrated for a sample of species, making comparisons with indices generated from GAMs, and introducing new methods of exploring covariate dependence.

Although we present and illustrate the work in terms of butterflies, it may be applied to other insect species, possibly after modification appropriate to their ecology. For example, the flightless longhorn beetle, *Dorcadion fuliginator*, takes two years to reach maturity (Baur et al. 2005), as do many dragonflies and some crickets. The models may also be adapted for the study of migrant bird populations.

## 2. MATERIALS AND METHODS

For any species, suppose counts of adults are recorded at  $S$  sites, each visited on up to  $T$  occasions, in each of  $Y$  successive years. Each can be treated as the realisation of a random variable from a suitable discrete distribution. For example, if this is taken as Poisson, with expectation  $\lambda_{i,j,k}$  for site  $i$ , visit  $j$  and year  $k$ , the likelihood has the form

$$L(\boldsymbol{\rho}, \boldsymbol{\mu}, \boldsymbol{\sigma}, N_1; \mathbf{y}) = \prod_{i=1}^S \prod_{j=1}^T \prod_{k=1}^Y \frac{\exp(-\lambda_{i,j,k}) \lambda_{i,j,k}^{y_{i,j,k}}}{y_{i,j,k}!},$$

where  $\{y_{i,j,k}\}$  are the counts and  $\rho$ ,  $\mu$ ,  $\sigma$  and  $N_1$ , are the model parameters which we describe in the next sections.

We adopt the Poisson distribution throughout, but there are other possibilities, such as negative-binomial or zero-inflated Poisson, for which an approximate concentrated likelihood approach is possible (Dennis et al. 2014). Alternatively, Pagel et al. (2014) accounted for overdispersion with respect to this simple model by a mixed log-normal-Poisson distribution.

The methods of the paper provide joint modelling of data obtained at different temporal scales. We consider two model types which are structurally different: a phenomenological model based simply on normal probability density functions and mechanistic models that are based upon stopover models, which involve mechanisms allowing for estimation of survival.

## 2.1. PHENOMENOLOGICAL MODEL FOR UNIVOLTINE SPECIES

For a univoltine species, the counts within a season increase from zero and then decrease to zero corresponding to the emergence and death of adult butterflies. This variation may be described by Normal probability density  $N(\mu_{i,k}, \sigma_{i,k}^2)$ , corresponding to site  $i$  and year  $k$ , so that for the  $j$ th visit at time  $t_{i,j,k}$  (e.g. week number in the season) we have

$$\lambda_{i,j,k} = N_{i,k} \frac{1}{\sigma_{i,k} \sqrt{2\pi}} \exp \left\{ -\frac{(t_{i,j,k} - \mu_{i,k})^2}{2\sigma_{i,k}^2} \right\}, \quad (1)$$

which we write as  $\lambda_{i,j,k} = N_{i,k} a_{i,j,k}$ , where  $N_{i,k}$  provides an estimate of relative abundance for site  $i$  in a given year,  $k$ , and  $\{a_{i,j,k}\}$  describes the seasonal variation over visits within that year. Thus for site  $i$  and year  $k$ , the counts for any visit have a Poisson distribution with mean value proportional to the Normal probability density function centred on  $\mu_{i,k}$ .

We allow the relative abundance  $N_{i,k+1}$ , for site  $i$  and year  $k+1$ , to depend upon that in the previous year,  $N_{i,k}$ , in a deterministic first-order auto-regressive manner via a growth rate,  $\rho_{i,k}$  which, assuming the species does not overwinter as an adult (in which case a model for multiple generations is required), we define as ‘‘productivity’’, i.e.  $N_{i,k+1} = \rho_{i,k} N_{i,k}$ . Developing this recursion over time provides

$$\lambda_{i,j,1} = N_{i,1} a_{i,j,1} \quad \text{and} \quad \lambda_{i,j,k} = N_{i,k} a_{i,j,k} = \left( N_{i,1} \prod_{m=1}^{k-1} \rho_{i,m} \right) a_{i,j,k} \quad \text{for } k > 1, \quad (2)$$

which is similar to the model in Freeman and Newson (2008), but with a seasonal component. The productivities,  $\{\rho_{i,k}\}$ , describe the successes of a given generation over sites ( $i$ ) for each year ( $k$ ) and represent products of the number of eggs laid per adult and the probability of each egg reaching the adult stage in the next generation. The expressions of Eq. (2) characterise the univoltine models of the paper, with different formulations for the seasonal pattern,  $\{a_{i,j,k}\}$ , providing different models, as we shall see for a mechanistic model formulation in Sect. 2.3.

## 2.2. PHENOMENOLOGICAL MODEL FOR BIVOLTINE SPECIES

Bivoltine butterfly species have two broods each year, with the adults of the second brood arising from the eggs laid by the adults of the first. We may extend the model above to describe counts from two annual broods by incorporating two Normal distributions in the model for  $\lambda_{i,j,k}$ . Thus we set

$$\lambda_{i,j,k} = N_{i,k,1} \frac{1}{\sigma_{i,k,1} \sqrt{2\pi}} \exp \left\{ -\frac{(t_{i,j,k} - \mu_{i,k,1})^2}{2\sigma_{i,k,1}^2} \right\} + N_{i,k,2} \frac{1}{\sigma_{i,k,2} \sqrt{2\pi}} \exp \left\{ -\frac{(t_{i,j,k} - \mu_{i,k,2})^2}{2\sigma_{i,k,2}^2} \right\},$$

which we may write as

$$\lambda_{i,j,k} \equiv N_{i,k,1} a_{i,j,k,1} + N_{i,k,2} a_{i,j,k,2},$$

where at site  $i$  in year  $k$  the relative abundance for the first brood is given by  $N_{i,k,1}$  and for the second brood by  $N_{i,k,2}$ . For the means and variances of the two Normal densities, the final subscripts designate brood, and we have  $\mu_{i,k,2} > \mu_{i,k,1}$ .

Whereas in [Dennis et al. \(2014\)](#) two broods are described by a mixture of probability density functions, here the relative abundance of a second brood in each year is assumed to depend on that of the first brood that year. Dependence between the two broods in any year is introduced by defining

$$N_{i,k,2} = \rho_{i,k,1} N_{i,k,1},$$

in addition to the between-year dependence, now given by

$$N_{i,k+1,1} = \rho_{i,k,2} N_{i,k,2}.$$

Thus,  $\rho_{i,k,1}$  represents the productivity of the first brood in a given year  $k$ , and  $\rho_{i,k,2}$  represents the productivity of the second brood, which feeds into the relative abundance of the first brood of the following year,  $N_{i,k+1,1}$ . So developing the recursion over time we write

$$\begin{aligned} \lambda_{i,j,1} &= N_{i,1,1} a_{i,j,1,1} + N_{i,1,2} a_{i,j,1,2} \\ &= N_{i,1,1} (a_{i,j,1,1} + \rho_{i,1,1} a_{i,j,1,2}), \end{aligned} \quad (3)$$

and

$$\begin{aligned} \lambda_{i,j,k} &= N_{i,k,1} a_{i,j,k,1} + N_{i,k,2} a_{i,j,k,2} \\ &= \left( N_{i,1,1} \prod_{m=1}^{k-1} \prod_{b=1}^2 \rho_{i,m,b} \right) a_{i,j,k,1} + \left( N_{i,1,1} \rho_{i,k,1} \prod_{m=1}^{k-1} \prod_{b=1}^2 \rho_{i,m,b} \right) a_{i,j,k,2} \\ &= N_{i,1,1} (a_{i,j,k,1} + \rho_{i,k,1} a_{i,j,k,2}) \prod_{m=1}^{k-1} \prod_{b=1}^2 \rho_{i,m,b} \quad \text{for } k > 1. \end{aligned} \quad (4)$$

The extension to multivoltine species with greater than two broods each year is immediate, though not greatly applicable to the UK species. The new development for bivoltine/multivoltine species is naturally based on the fact that the relative size of a given brood depends on the productivity of the previous brood. Notationally, we denote the phenomenological models by  $P_B$ , where  $B$  is the number of broods.

### 2.3. MECHANISTIC AND STOPOVER MODELS

Relatively little is known regarding butterfly survival, and what is known results from local short-term mark-recapture programmes, which are expensive. To build survival into our models we introduce additional parameters, the emergence times of adults, which are typically unknown and of interest in their own right as indicators of phenological change in a specific, key point in a species' life-cycle. We do this as follows.

Suppose first of all that there is only one brood, and a site abundance  $N_{i,k}$  for site  $i$  and year  $k$ . In order to describe the emergence times we introduce parameters  $\beta_{i,d-1,k}$ , which describe the proportions of  $N_{i,k}$  emerging at site  $i$  and just prior to visit  $d$  in year  $k$ . The expected number of individuals at site  $i$  at time  $t_{i,j,k}$  in year  $k$  is given as

$$\lambda_{i,j,k} = N_{i,k} a_{i,j,k} = N_{i,k} \left\{ \sum_{d=1}^j \beta_{i,d-1,k} \left( \prod_{m=d}^{j-1} \phi_{i,m,k} \right) \right\}, \quad (5)$$

where the index  $d = 1, \dots, j$  indicates the possible times of emergence for an individual detected on visit  $j$ . The parameters  $\beta_{i,d-1,k}$  describe the proportions of  $N_{i,k}$  emerging at site  $i$  and visit  $d$  in year  $k$ , such that  $\sum_{d=1}^T \beta_{i,d-1,k} = 1$ , for each site  $i$  and year  $k$ . We define  $\phi_{i,m,k}$  as the probability that an individual that is present at site  $i$  at visit  $m$  in year  $k$ , will remain at that site until visit  $m + 1$ . So for example,  $\lambda_{i,3,k} = N_{i,k} (\beta_{i,0,k} \phi_{i,1,k} \phi_{i,2,k} + \beta_{i,1,k} \phi_{i,2,k} + \beta_{i,2,k})$ .

In order that the emergence parameters have the right type of shape we can set

$$\beta_{i,d-1,k} = F_{i,k}(t_{i,d,k}) - F_{i,k}(t_{i,d,k} - 1), \quad (6)$$

where  $F_{i,k}(t_{i,d,k}) = P(X \leq t_{i,d,k})$  for  $X \sim N(\mu_{i,k}, \sigma_{i,k}^2)$ , where  $\mu_{i,k}$  is the mean date of emergence and  $\sigma_{i,k}^2$  is the associated variance. For each  $i, k$ ,  $\beta_{i,0,k} = F_{i,k}(1)$  and  $\beta_{i,T-1,k} = 1 - F_{i,k}(T - 1)$ .

This is a simple stopover model, proposed for butterfly data by [Matechou et al. \(2014\)](#); see also [Dennis et al. \(2014\)](#). Stopover models are used in describing data on migrating birds, which rest and feed during their journey at particular stopover sites where observations take place. Typically, the resulting counts, graphed over time, reflect successive waves of birds arriving, staying and then leaving. [Matechou et al. \(2014\)](#) observe that this is the same pattern seen when adult butterflies are counted within a season, with for example bivoltine species analogous to two waves of birds observed at a stopover site.

In [Matechou et al. \(2014\)](#), the above model is extended to account for multivoltine data, and the expression of Eq. (6) then becomes a mixture of terms, each of which is an area under an appropriate probability density function. In the multivoltine case, we need a

different dynamic mechanistic model, in order to allow for the abundance of one brood to feed into that of a succeeding brood, during the same year.

In the univoltine dynamic stopover model, the recursions of Eq. (2) apply, but now with the different specification of  $\{a_{i,j,k}\}$  provided by Eq. (5). For the multivoltine case in the dynamic mechanistic model, we assign a separate site abundance to each brood in a year. Thus, we assume the two broods for bivoltine species to be separate such that, for site  $i$ , visit  $j$  and brood  $b$ , in year  $k$ , we extend Eq. (5) to give

$$a_{i,j,k,b} = \sum_{d=1}^j \beta_{i,d-1,k,b} \left( \prod_{m=d}^{j-1} \phi_{i,d,k,b} \right), \quad \text{for } b = 1, 2, \quad (7)$$

where we define  $\{\phi_{i,d,k,b}\}$  as the appropriate survival probabilities of an individual from one visit to the next, which are now estimated separately for each brood. This development is an extension of that in the original specification of Matechou et al. (2014). For brood  $b$ , the parameters  $\{\beta_{i,d-1,k,b}\}$  describe the proportions of  $N_{i,k,b}$  arriving at visit  $d$ , and are modelled here using Normal distributions, so that

$$\beta_{i,d-1,k,b} = F_{i,k,b}(t_{i,d,k}) - F_{i,k,b}(t_{i,d,k} - 1),$$

where  $F_{i,k,b}(t_{i,d,k}) = \Pr(X \leq t_{i,d,k})$ , for  $X \sim N(\mu_{i,k,b}, \sigma_{i,k,b}^2)$ , and  $\mu_{i,k,b}$  is the appropriate mean date of emergence of adults for brood  $b$  and  $\sigma_{i,k,b}^2$  is the corresponding variance. For each  $i, k$ , and  $b$ ,  $\beta_{i,0,k,b} = F_{i,k,b}(1)$  and  $\beta_{i,T-1,k,b} = 1 - F_{i,k,b}(T-1)$ . The recursions of Eqs. (3) and (4) then apply, but now with the new specification of  $\{a_{i,j,k,b}\}$  from Eq. (7). Notationally, we specify the dynamic mechanistic model by  $M_B$ , where  $B$  is the number of broods.

#### 2.4. CONCENTRATED LIKELIHOOD

We fit models to data by maximum likelihood. As in Dennis et al. (2014), the number of parameters in the likelihood can be reduced by  $S$ , using a concentrated likelihood approach.  $S$  is typically large for these models and so computational efficiency is substantially increased. We consider first the univoltine case. Using Eq. (2), apart from an additive constant, the log-likelihood for site  $i$  may be written as

$$\begin{aligned} \ell_i = \text{Log}(L_i) = & \sum_{j=1}^T \left[ -N_{i,1} a_{i,j,1} + y_{i,j,1} \log(N_{i,1} a_{i,j,1}) \right. \\ & \left. + \sum_{k=2}^Y \left\{ -N_{i,1} a_{i,j,k} \prod_{m=1}^{k-1} \rho_{i,m} + y_{i,j,k} \log \left( N_{i,1} a_{i,j,k} \prod_{m=1}^{k-1} \rho_{i,m} \right) \right\} \right]. \quad (8) \end{aligned}$$

For the data from all sites, the log-likelihood is  $\ell = \sum_{i=1}^S \ell_i$ . Using Eq. (8) we obtain

$$\frac{\partial \ell}{\partial N_{i,1}} = \sum_{j=1}^T \left\{ -a_{i,j,1} + \frac{y_{i,j,1}}{N_{i,1}} + \sum_{k=2}^Y \left( -a_{i,j,k} \prod_{m=1}^{k-1} \rho_{i,m} + \frac{y_{i,j,k}}{N_{i,1}} \right) \right\},$$

and equating to zero we find

$$N_{i,1} = \sum_{j=1}^T \frac{\sum_{k=1}^Y y_{i,j,k}}{a_{i,j,1} + \sum_{k=2}^Y a_{i,j,k} \prod_{m=1}^{k-1} \rho_{i,m}}. \quad (9)$$

We note how  $N_{i,1}$  is a weighted sum over visits of totals at site  $i$  across years. Thus, despite an apparent strong dependence of  $\{N_{i,k}\}$  on  $\{N_{i,1}\}$  in (2), this is only a consequence of the deterministic links between the  $\{N_{i,k}\}$ , and all data contribute to the estimation of  $\{N_{i,1}\}$ , and hence  $\{N_{i,k}\}$ . Substitution of the expressions for  $\{N_{i,1}\}$  from (9) in (8) results in a concentrated likelihood, which is maximised with respect to only the parameters associated with  $\boldsymbol{\rho}$  and  $\mathbf{a}$  (which contain the elements of  $\boldsymbol{\mu}$  and  $\boldsymbol{\sigma}$ ). Estimation of  $\{N_{i,1}\}$  is then made by substituting estimates of  $\{a_{i,j,k}\}$  and  $\{\rho_{i,m}\}$  into (9). The above approach holds for both phenomenological and mechanistic models.

The concentrated likelihood for the bivoltine case is given similarly in the appendix. We maximise the concentrated likelihoods using the `optim` function in R (R Core Team 2015), with the limited-memory BFGS algorithm (Byrd et al. 1995). Associated R code for the dynamic models is provided in the Supplementary Material.

## 2.5. ANNUAL INDEX OF ABUNDANCE

In the following, we write  $\hat{\theta}$  for the maximum-likelihood estimate of  $\theta$ , for any parameter  $\theta$ . The averages of the relative site abundance estimates, for each year  $k$ , are used to create an index of abundance  $G_k$  for year  $k$ . For a univoltine species, we set

$$G_k = \frac{1}{S} \sum_{i=1}^S \hat{N}_{i,k} = \begin{cases} \frac{1}{S} \sum_{i=1}^S \hat{N}_{i,1} & \text{if } k = 1 \\ \frac{1}{S} \sum_{i=1}^S \left( \hat{N}_{i,1} \prod_{m=1}^{k-1} \hat{\rho}_{i,m} \right) & \text{if } k > 1, \end{cases} \quad (10)$$

for  $k = 1, \dots, Y$ , from equations (2). Similarly for the bivoltine case, we estimate an index  $G_{k,b}$  for each brood,  $b = 1, 2$ , as

$$G_{k,1} = \frac{1}{S} \sum_{i=1}^S \hat{N}_{i,k,1} = \begin{cases} \frac{1}{S} \sum_{i=1}^S \hat{N}_{i,1,1} & \text{if } k = 1 \\ \frac{1}{S} \sum_{i=1}^S \left( \hat{N}_{i,1,1} \prod_{m=1}^{k-1} \prod_{b=1}^2 \hat{\rho}_{i,m,b} \right) & \text{if } k > 1, \end{cases} \quad (11)$$

and

$$G_{k,2} = \frac{1}{S} \sum_{i=1}^S \hat{N}_{i,k,2} = \begin{cases} \frac{1}{S} \sum_{i=1}^S \hat{N}_{i,1,1} \hat{\rho}_{i,k,1} & \text{if } k = 1 \\ \frac{1}{S} \sum_{i=1}^S \left( \hat{N}_{i,1,1} \hat{\rho}_{i,k,1} \prod_{m=1}^{k-1} \prod_{b=1}^2 \hat{\rho}_{i,m,b} \right) & \text{if } k > 1, \end{cases} \quad (12)$$



for  $k = 1, \dots, Y$ , making use of the recursions demonstrated in Eqs. (3) and (4). The separate brood indices can be added to produce a single, annual index but there is potentially great ecological benefit in maintaining them separately, as each corresponds to different times of year and may be driven by different environmental factors.

Once indices are formed they are plotted against year, and we shall see examples in Sect. 3. Standard errors for the indices can be obtained via bootstrapping, as for other methods (Dennis et al. 2013; Dennis et al. 2014). Error bars are not presented here for clarity, but in general the differences between the indices derived from the dynamic models and alternative methods (which we explore in the next section) are smaller than the size of the errors.

## 2.6. APPLICATION

We apply the dynamic models to national monitoring scheme data for a subset of UK butterfly species. The UK Butterfly Monitoring Scheme (UKBMS) is the primary source of count data for UK butterflies. The scheme relies on recorders who count butterflies under favourable conditions each week between early April and late September, the main period for butterfly activity. This results in a maximum of  $T = 26$  each year, though typically not all of the 26 designated visits are made, so the data do not need to be equally spaced. The UKBMS has grown gradually since it began in 1976 to over 1100 sites monitored in 2012 (Botham et al. 2013). Population trends are typically calculated annually for 56 of the 59 butterfly species regularly found in the UK.

Many studies of UKBMS data involve application to a single illustrative species (Matechou et al. 2014; Pagel et al. 2014). We demonstrate the dynamic models with application to a sample of representative, taxonomically and ecologically diverse species. Six univoltine and five bivoltine species were selected, with varying range size, habitat requirements and phenologies, although very scarce, habitat-specialist species, which generally have limited data, were not considered in this analysis. Each model was fitted to data for 1978–2011. The UK butterfly transect data used in this study are archived by the UKBMS (<http://www.ukbms.org>).

Sites at which the species of interest was never recorded or at which monitoring was undertaken for fewer than five years were excluded from this analysis. For illustration, a subset of 100 monitored sites was randomly selected for each species, with the exception of Holly Blue, for which a sample of up to 200 sites was instead taken, since using only 100 sites produced bias in the estimates of productivity.

We illustrate the performance of the dynamic models in terms of abundance indices, productivity, survival and phenology. Additional figures and tables are given in Appendix S1 of the Supplementary Material. This will be done for the samples of the univoltine and bivoltine species, with and without the addition of covariates. Where parameters were assumed to be constant spatially the subscript for site,  $i$ , is omitted. In models for bivoltine species, we let  $\mu_2 = \mu_1 + \mu_d$ , where  $\mu_1 \geq 0$  and  $\mu_d > 0$ , to ensure that  $\mu_2 > \mu_1$ .

The covariates we select are northing and measures of temperature. They were chosen to demonstrate the potential of the models, and may not be optimal. All covariates were standardised to have zero mean and unit variance. We use monthly mean and minimum Central England Temperature data (Parker et al. 1992).

The average minimum daily temperature during October–March was used as a covariate for overwinter productivity. For bivoltine species, the mean temperature within the flight period of the first brood was used to describe productivity of the first brood. Productivities, which are necessarily positive, were regressed on the log scale. Survival in mechanistic models was logistically regressed on mean temperature within the flight period of the brood of interest. Scientific names and approximate flight periods for the species studied are provided in Table S1.1, and the latter were used to indicate the relevant temperature covariates. Due to interest in the possible effect of covariates on estimates of survival, we primarily use mechanistic models when covariates are employed and phenomenological models otherwise, however alternatives are also possible.

### 3. RESULTS

#### 3.1. INDICES

Indices of abundance are derived from estimates of annual productivity and estimates of initial abundance from the dynamic model, as described in Sect. 2.5. Here  $\mu$  and  $\sigma$  have been considered constant, although varying these between years provides useful information and we shall see examples of this later, but it had no distinguishable effect on indices of abundance. We compare relative abundance indices for model  $P_1$  and an approach with GAM-based models for seasonal patterns, currently adopted by the UKBMS and described by Dennis et al. (2013). To compare the different indices, each index was standardised to have zero mean and unit variance. An additional comparison with the generalised abundance index (GAI) approach (Dennis et al. 2014) and consideration of goodness-of-fit are given in Appendix S2 of the Supplementary Material.

The fitted phenomenological dynamic models discussed in the context of indices have 35 and 71 parameters for  $B = 1, 2$ , respectively. Given that  $Y = 34$ , in the univoltine case there are 33 annual estimates  $\rho_k$ , as well as  $\mu$  and  $\sigma$ , and in the bivoltine case, there are 34 parameters  $\rho_{k,1}$  and 33 parameters  $\rho_{k,2}$ , in addition to  $\mu_1, \mu_d, \sigma_1$  and  $\sigma_2$ .

Figure 1a gives a comparison between annual indices of abundance for six univoltine species. There is a good agreement between the indices resulting from the dynamic model and the standard GAM approach (Dennis et al. 2013). By estimating an index for each brood (Eqs. 11 and 12), dynamic models  $P_2$  allow us to add more information to indices for bivoltine species, which we illustrate in two different ways in Figs. 2a and S1.1. We see how the dynamic model allows us to elaborate the indices produced by the GAM approach, by providing a separate index for each brood in the bivoltine case. This could, for example, reveal differing trends between broods.

For model verification, Appendix S3 of the Supplementary Material summarises the results of applying the dynamic models to simulated data.

#### 3.2. PRODUCTIVITY

Figure 1b presents estimates of productivity for the univoltine species, from fitting model  $P_1$ . Values of  $\rho_k$  greater than unity indicate growth compared to the previous year, and values

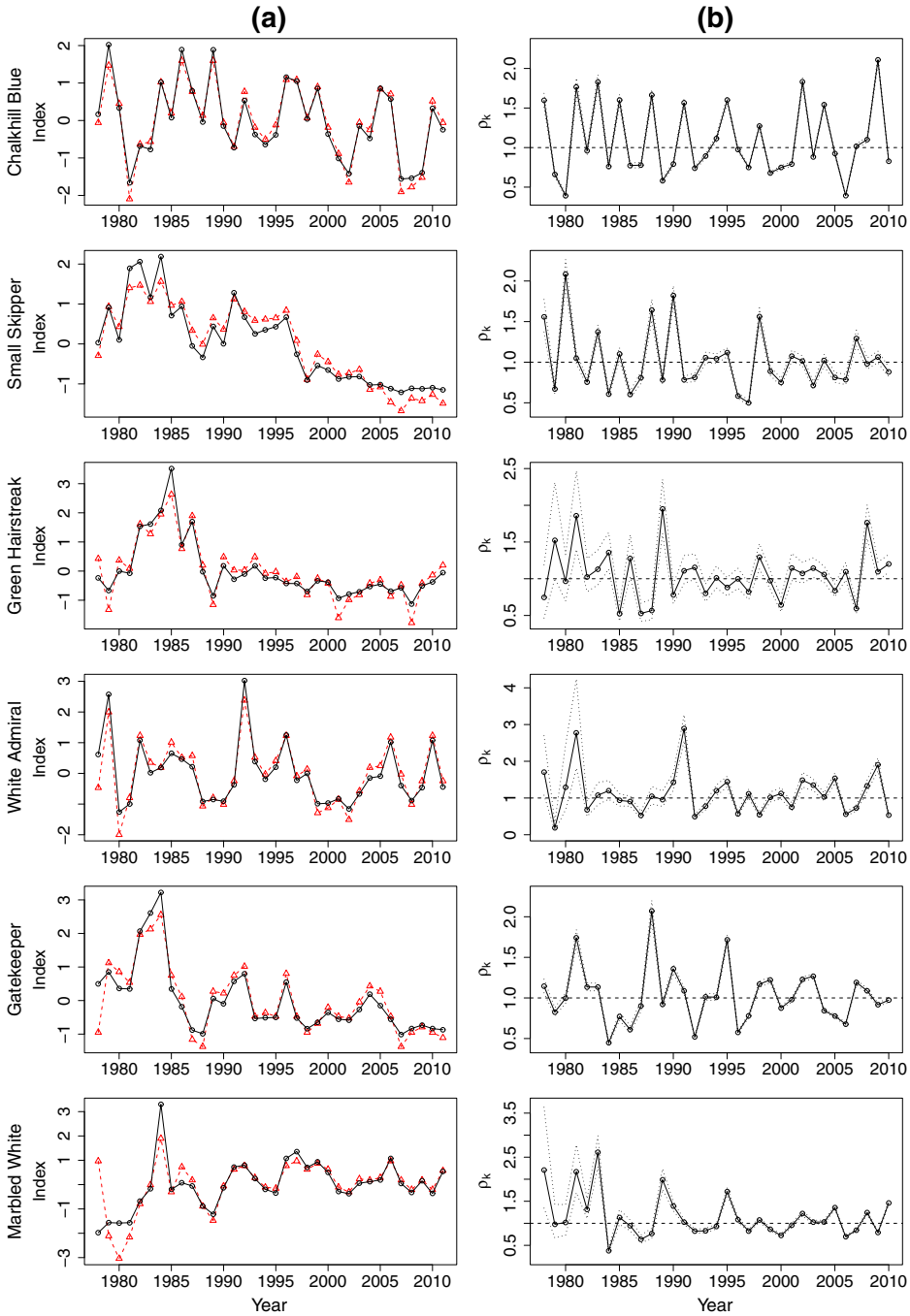


Figure 1. **a** Relative abundance indices for each univoltine species from model  $P_1$  (black, open circles, solid lines) and the GAM approach (red, open triangles, dashed lines) and **b** annual estimates of productivity,  $\rho_k$ , from model  $P_1$ , which was fitted to estimate  $\rho_k$  across sites for each year. The horizontal dashed line separates productivities above/below unity, corresponding to growth/decline compared to the previous year. Dotted lines represent 95 % confidence intervals for productivity (Color figure online).

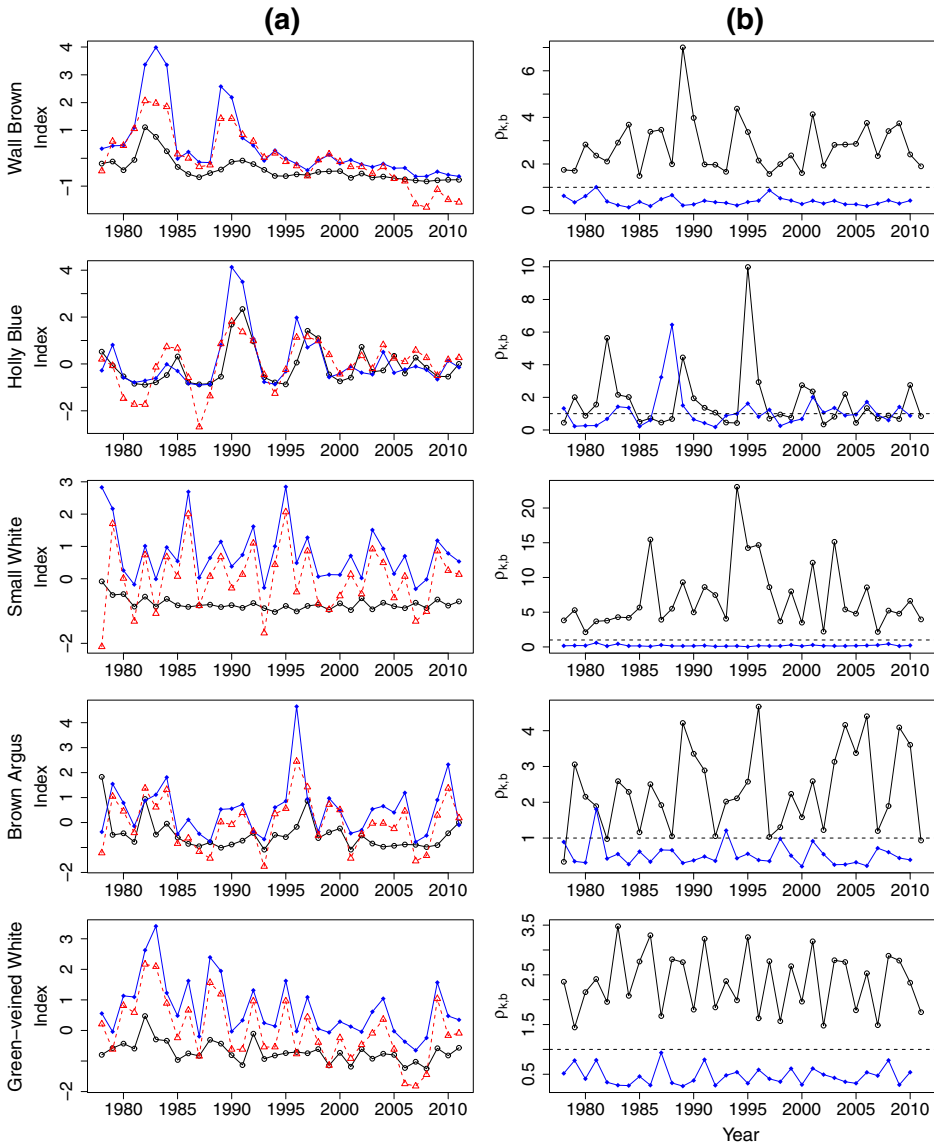


Figure 2. **a** Relative abundance indices for each bivoltine species for the first (*black, open circles, solid lines*) and second (*blue, closed diamonds, solid lines*) broods from model  $P_2$  and the GAM approach (*red, open triangles, dashed lines*) and **b** annual estimates of productivity for the first ( $\rho_{k,1}$ , *black, open circles*) and second ( $\rho_{k,2}$ , *blue, closed diamonds*) brood from model  $P_2$ , which was fitted to estimate  $\rho_{k,b}$  across sites for each brood and year. The horizontal dashed line separates productivities above/below unity, corresponding to growth/decline compared to the previous brood (Color figure online).

less than unity indicate decline. Hence as anticipated we see a tendency for productivities less than unity for species in decline, such as Small Skipper in recent years, while for Marbled White productivities tend to be above unity during the initial period of growth, followed by recent fluctuations about unity, when the population appears to be relatively stable.

Figure 2b presents estimated productivities for each brood for five bivoltine species, using model  $P_2$ . Values above/below unity represent growth/decline relative to the previous

brood. In Figure S1.2, we see how the productivities reflect the relative sizes of the fitted seasonal curves, for which the average over the series is shown (standardised to sum to unity). The relative sizes of the broods will actually vary with productivity each year.

Figure 3 shows the results of including covariates in the dynamic models, in this case for model  $M_1$  with  $\rho_{i,k}$  and  $\phi_{i,k}$  (which we revisit in Sect. 3.3) varying with temperature and northing. It is interesting that with the exception of Gatekeeper, higher productivity is significantly associated with cooler winters and in all cases with more Northerly latitudes (regression coefficients and associated standard errors are presented in Table S1.2a).

Figure 4 shows the effect of adding covariates for bivoltine species, in this case for model  $M_2$  with productivity varying with temperature and northing, and survival varying with temperature, which we discuss further in Sect. 3.3. As detailed in Sect. 2.6, first-brood productivity was associated with the mean temperature during the first brood, and second-brood productivity with the minimum winter overwinter temperature. Associations of first-brood productivity,  $\rho_{i,k,1}$ , with northing and weather varied between species, and regression coefficients for the slope parameters were generally significant (Table S1.3a). The association of higher productivity with cooler winters shown for univoltine species is also found for the second brood of the bivoltine species, with the exception of Wall Brown and Holly Blue, which is a common garden visitor, unlike the other species which favour grasslands, as well as gardens in the case of Small White and Green-veined White.

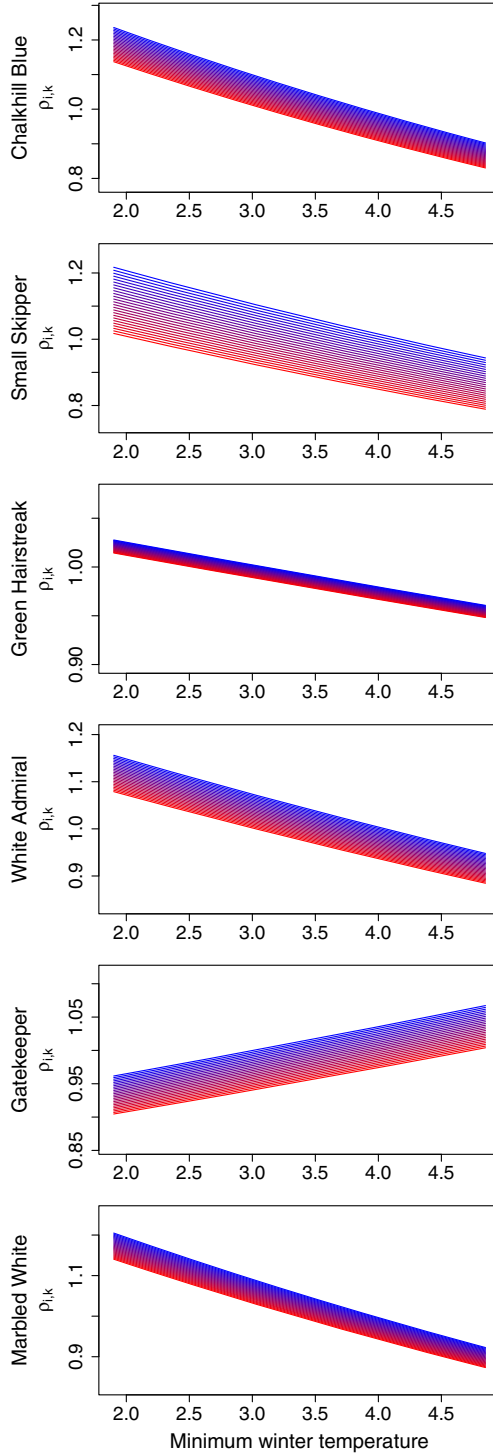
Given an estimate of productivity for each year, if desired the geometric mean of the productivities over time may be used to provide a simple comparison between species.

### 3.3. SURVIVAL

The mechanistic models allow estimation of the survival probabilities,  $\phi$ , of butterflies, from which adult life expectancies (in weeks) can be estimated by  $1/(1 - \phi)$ , assuming that a species does not overwinter as an adult. Variation in life expectancy with temperature is displayed for univoltine species in Figure S1.3, and for bivoltine species in Figure S1.4, based on the models fitted with covariates in the previous section. Tables S1.2 and S1.3 show the parameter estimates and associated standard errors from the  $M_B$  models with covariates. For comparison, estimates are also included for the  $P_1$  and  $P_2$  models with covariates for  $\rho$ , which are not presented in the figures, but produce analogous estimates of the shared parameters. There are differences in  $\mu$  and  $\sigma$  since in the mechanistic model  $\mu$  represents the mean date of emergence which will be earlier than the mean flight date, and  $\sigma$  relates to the length of the period of emergence, which will be shorter than the length of the flight period in the  $P_1$  model. The associated errors for  $\mu$  and  $\sigma$  are smaller for the  $P_1$  than for the  $M_1$  model. For the bivoltine species, there is more variation in the estimates from  $P_2$  and  $M_2$ . As in the univoltine case, standard errors from the phenomenological model tend to be smaller than those from the mechanistic model. The  $M_B$  models with covariates have 8 and 14 parameters for  $B = 1, 2$ , respectively, compared to the  $P_B$  models with 5 and 10 parameters for  $B = 1, 2$ , respectively. In these cases, reduced precision is a consequence of greater model complexity.

For univoltine species, there was a significant negative association of life expectancy with higher average temperature during the flight period for four out of six species (Figure S1.3;

Figure 3. Predicted productivity with varying minimum overwinter temperature from model  $M_1$ . Each line represents one of 25 equally spaced northing values within the species range (*red* at southern sites and *blue* at northern sites). Model  $M_1$  was fitted with productivity,  $\rho_{i,k}$ , and survival probability,  $\phi_{i,k}$ , regressed on temperature and northing (Color figure online).



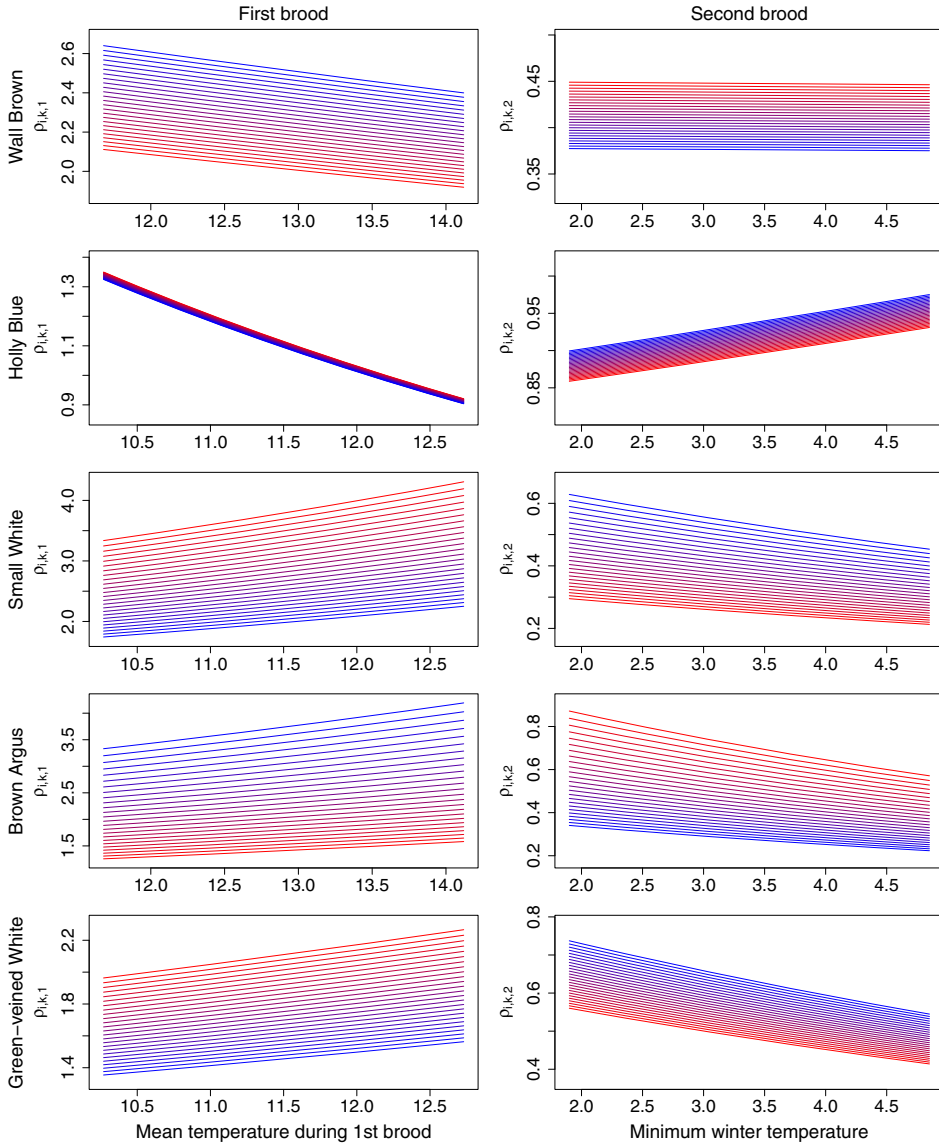


Figure 4. Predicted productivity with varying temperature from model  $M_2$ . The mean temperature during the first brood, and the minimum overwinter temperature were used as covariates for productivity of the first and second brood, respectively. Each *line* represents one of 25 equally spaced Northing values within the species range (red at southern sites and blue at northern sites). Model  $M_2$  was fitted with productivity,  $\rho_{i,k,b}$ , for each brood,  $b$ , regressed on temperature and northing and survival probability,  $\phi_{i,k,b}$ , for each brood,  $b$ , regressed on temperature (Color figure online).

Table S1.2a). Four univoltine species indicated significantly greater survival at southerly sites. Standard errors in Table S1.2a) are generally small, but are large for two instances for Green Hairstreak, which exhibit flatness in the associated plots (Figs. 3, S1.3).

As for the associations of first-brood productivity with weather, in bivoltine species we find that the variation in first brood life expectancy with temperature differs between the species sampled, and slope estimates were only significant for three out of five species

(Table S1.3a). With the exception of Holly Blue, life expectancy for the second brood of the bivoltine species increased significantly with temperature. Fitting model  $M_2$  with covariates for northing and weather on  $\rho$  and  $\phi$  for each brood produced unrealistic estimates of lifespan for Brown Argus and Holly Blue, hence in Figure S1.4 we allow  $\phi$  in the  $M_2$  model to vary with temperature and brood only. This requires further investigation, but is likely to be due to the relatively large number of parameters in model  $M_2$  and/or relatively small size of the sample. The corresponding standard errors for this model are sometimes large, for example, for Holly Blue (Table S1.3a).

### 3.4. PHENOLOGY

Here we demonstrate the potential to produce estimates of phenology using the dynamic models. The  $P_1$  and  $P_2$  models were fitted with  $\rho$ ,  $\mu$  and  $\sigma$  each varying with year. Hence the  $P_1$  model requires 101 parameters to be estimated, corresponding to 33 parameters for  $\rho_k$  and 34 parameters each for  $\mu_k$  and  $\sigma_k$ . Similarly the  $P_2$  model has 203 parameters: 34 for  $\rho_{k,1}$ , 33 for  $\rho_{k,2}$ , and 34 each for  $\mu_{k,1}$ ,  $\mu_{k,d}$ ,  $\sigma_{k,1}$  and  $\sigma_{k,2}$ . To identify potential phenological trends, the models were also fitted with the parameters of interest regressed upon year (indicated by blue lines), as in the models fitted to univoltine species for comparison with the GAI in Appendix S2 of the Supplementary Material. We perform simple linear regressions post model-fitting to identify potential trends between  $\mu$  and productivity  $\rho$ , where green lines indicate significant regressions ( $p$ -value  $< 0.05$ ).

Figure 5 gives the mean and standard deviation of the flight periods for the univoltine species and corresponding figures for the bivoltine species are given in Figures S1.5 and S1.6. Figures 5a and S1.5 suggest that the mean flight period date,  $\mu$ , has advanced for all species and broods, which is consistent with what is expected under climate change (Sparks and Yates 1997; Roy and Sparks 2000). From Figures 5b and S1.6 we see that the length of the flight period has generally increased, also in agreement with previous findings (Roy and Sparks 2000). Table S2.1 suggests significant increases in  $\sigma$  for 5 out of 6 univoltine species. The location of the fitted line for the Marbled White  $\sigma_k$  in Figure 5b is due to the increase in sample size over time giving more weight to the later years. Figures S1.5 and S1.6 show a small number of outliers which require further investigation.

With the exception of Green Hairstreak, for the six univoltine species there was no clear relationship between  $\mu_k$  and  $\rho_k$  (Figure S1.7). For Green Hairstreak, which emerges early in the season, lower productivities are associated with an earlier flight period, which may lead to declines if advances in phenology continue with changes in climate. For most of the five bivoltine species, significant patterns between the mean flight period for each generation and the associated productivity were not found (Figure S1.8). However for Brown Argus and Green-veined White, productivity of the second generation was lower when the mean flight period date of the second brood,  $\mu_{k,2}$ , was advanced.

These results show that the dynamic models predict phenological changes consistent with expected patterns. The dynamic models allow for improved estimates of phenology to be studied in combination with demographic parameters, to reveal potential novel insights. Changes in phenology may also be modelled using the mechanistic models, in order to separate changes in emergence time from changes in survival.



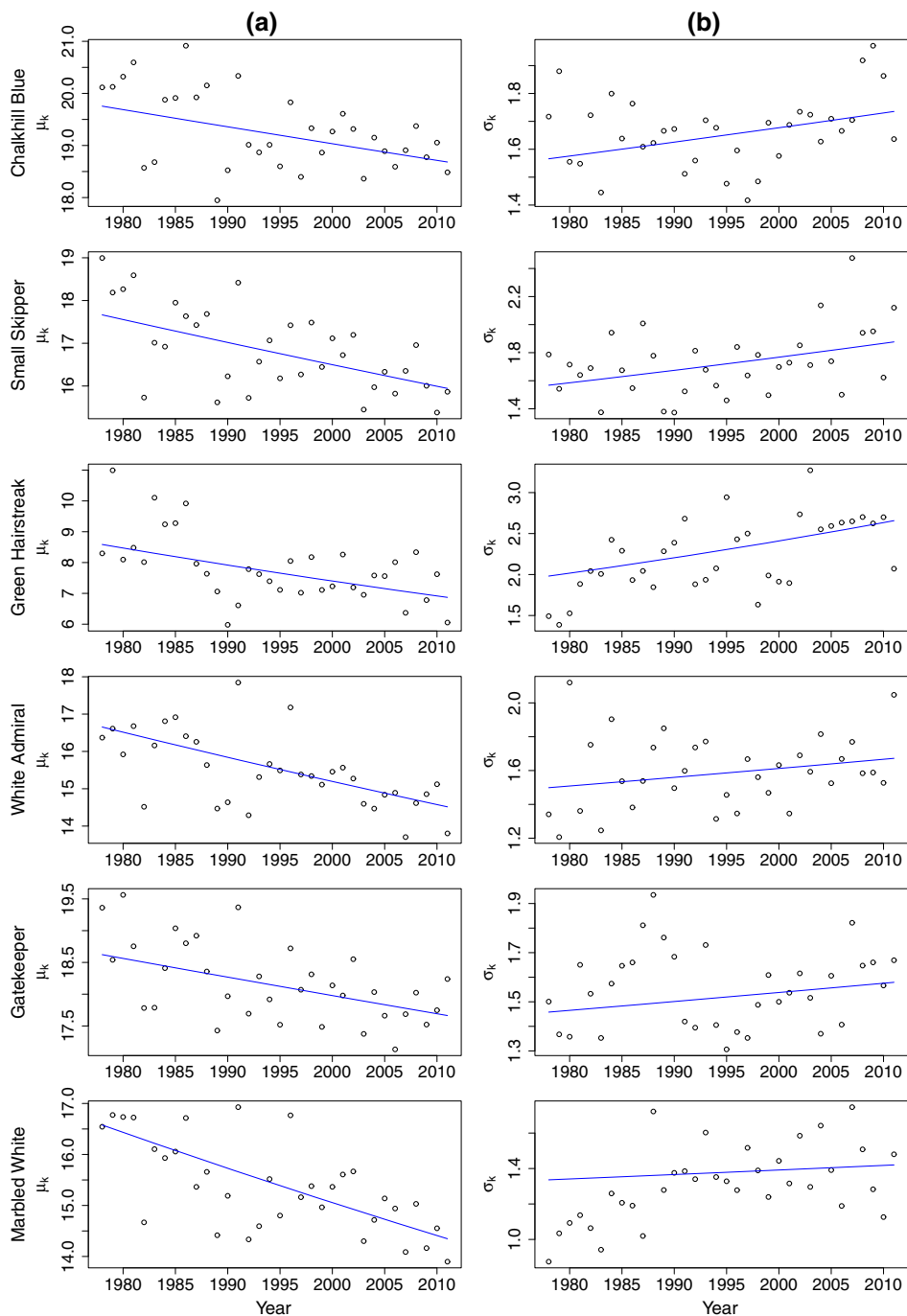


Figure 5. Annual estimates of **a**  $\mu_k$  and **b**  $\sigma_k$  from model  $P_1$ , which was fitted to estimate  $\rho_k$ ,  $\mu_k$  and  $\sigma_k$  across sites for each year. *Blue lines* indicate fitting log-linear regressions on year for  $\mu$  and  $\sigma$ , as in Table S2.1a) (Color figure online).

## 4. DISCUSSION

The dynamic model framework allows novel investigation of the drivers of fluctuations in abundance and provides a basis that can be adapted to both the study species and research aim. We have presented only a preliminary application. The methods of [Dennis et al. \(2014\)](#), which model data for each year separately, may be better suited for estimating indices of abundances efficiently (see Table S2.1), whereas dynamic models provide additional information of value for understanding demography. However, the agreement of the indices obtained from the different methods provides confidence that the dynamic models are performing correctly, and that for multivoltine species indices may be derived separately for each brood.

For the majority of the sample species, higher overwinter productivity was associated with cooler winters, which may act to reduce the impact of pathogens. Variability in lifespan and first brood productivity of bivoltine species differed more between species. Given that species have different life-histories, further research may look for trait-based variation, for example overwintering stage: egg, larva, chrysalis or adult, all of which may be affected most severely by different environmental factors. For example, [Diamond et al. \(2011\)](#) explored relationships between changes in date of first appearance and species' traits.

Further work is needed to explore the relevant covariates driving changes in productivity, survival and phenology. Spatial covariates such as habitat/land-cover may describe additional variation in the parameters. Inclusion of local weather could identify the period within the life-cycle for which weather has the most impact on the adult stage. Growing degree-days may also be explored ([Hodgson et al. 2011](#)). In this study, covariates were included additively on a logistic linear scale, whereas true relationships may be non-linear, for example productivity/survival might be limited by extremes in weather. The models could also be extended to model variation in productivity stochastically.

Alternatives to the Normal distribution for describing seasonal variation could be explored, for example to describe skewness ([Calabrese 2012](#)). This study has only accounted for species which are distinctly univoltine or bivoltine. A spline may be used to define complex seasonal patterns ([Dennis et al. 2014](#)), and the models could be extended to allow more than two broods each year. The models may be developed to accommodate variation in voltinism, where the first generation contributes to both the second generation within the same year and first generation the following year, with relevance for study of potential "lost generations" ([Dyck et al. 2015](#)). The dynamic models may also be used to study species which aestivate under hot summer conditions ([Spieth et al. 2011](#); [Grill et al. 2013](#)).

The dynamic models produce realistic estimates of parameters relevant to phenology, providing further validation of the models. Phenological studies have typically involved measures such as mean first encounter, mean peak encounter and mean length of the flight period ([Roy and Sparks 2000](#); [Diamond et al. 2014](#); [Karlsson 2014](#)), which may be driven by observer behaviour. The improved estimates of phenology from dynamic models provide the opportunity to study linkages between changes in phenology and changes in abundance and productivity, for example phenological mismatch ([Hindle et al. 2015](#)).

Using a phenomenological model may be optimal in scenarios with limited data, but the mechanistic model allows for additional insights by estimating survival. Spatio-temporal

variation in the lifespans of butterflies has had limited attention, as have potential linkages with other parameters, for example to explore how phenology affects survival, or whether variation in survival can influence productivity. Using a mechanistic model separates relevant parameters, for example to determine whether an increase in flight period length is due to an extended period of emergence, or increased lifespan. The model could be adapted to explore synchrony in populations (Powney et al. 2010), either between sites of a given species or across sites but between multiple species, by incorporating random effects (Lahoz-Monfort et al. 2011, 2013), for example in the  $\rho$  parameter for productivity. Density dependence, which has been highlighted for some butterflies (Nowicki et al. 2009), may be incorporated here in productivity and/or survival by introducing a dependency on the relative abundance. Additionally, allowing for spatial dependence of  $\rho$  and autocorrelation in abundance may be advantageous (Johnson et al. 2012). Pagel et al. (2014) included spatially autocorrelated random effects when modelling mean population density, but did not account for the within-year variability in counts.

For some threatened, conservation-priority UK butterflies, such as Large Blue *Phengaris arion*, Brown Hairstreak *Thecla betulae* and Marsh Fritillary *Euphydryas aurinia*, data are available on other stages of the butterfly life-cycle, such as counts of caterpillars or eggs. An attraction of the model framework proposed is the potential for the incorporation of data from multiple stages of the life-cycle, which could aid the monitoring and conservation of rarer species for which coverage from standard monitoring schemes can be limited.

The dynamic models may address the “lack of mechanistic understanding about factors driving butterfly population dynamics” (Isaac et al. 2011). Future application will generate hypotheses for further investigation, with the potential to illuminate features of butterfly phenology and demography which are at present poorly understood.

## ACKNOWLEDGEMENTS

This work was part-funded by EPSRC grants EP/1000917/1 and EP/P505577/1. We thank Martin Ridout, Marc Kéry and Michael Schaub for their useful comments. The UKBMS is operated by the Centre for Ecology & Hydrology and Butterfly Conservation and funded by a multi-agency consortium including the Countryside Council for Wales, Defra, the Joint Nature Conservation Committee, Forestry Commission, Natural England, the Natural Environment Research Council, and Scottish Natural Heritage. The UKBMS is indebted to all volunteers who contribute data to the scheme.

[Received January 2015. Accepted August 2015. Published Online August 2015.]

## 5. APPENDIX: CONCENTRATED LIKELIHOOD FOR BIVOLTINE SPECIES

Using Eq. (4), the log-likelihood for site  $i$  is given, apart from an additive constant, by

$$\ell_i = \text{Log}(L_i) = \sum_{j=1}^T \left[ -N_{i,1,1} (a_{i,j,1,1} + \rho_{i,1,1} a_{i,j,1,2}) + y_{i,j,1} \log \{N_{i,1,1} (a_{i,j,1,1} + \rho_{i,1,1} a_{i,j,1,2})\} \right]$$

$$\begin{aligned}
& + \sum_{k=2}^Y \left\{ -N_{i,1,1} (a_{i,j,k,1} + \rho_{i,k,1} a_{i,j,k,2}) \prod_{m=1}^{k-1} \prod_{b=1}^2 \rho_{i,m,b} \right. \\
& \left. + y_{i,j,k} \log \left( N_{i,1,1} (a_{i,j,k,1} + \rho_{i,k,1} a_{i,j,k,2}) \prod_{m=1}^{k-1} \prod_{b=1}^2 \rho_{i,m,b} \right) \right\}, \quad (13)
\end{aligned}$$

where we have defined  $\{a_{i,j,k,b}\}$  and  $\{\rho_{i,k,b}\}$ , for site  $i$ , visit  $j$  and brood  $b$  in year  $k$  in the main paper. This gives

$$\begin{aligned}
\frac{\partial \ell}{\partial N_{i,1,1}} &= \sum_{j=1}^T \left[ - (a_{i,j,1,1} + \rho_{i,1,1} a_{i,j,1,2}) + \frac{y_{i,j,1}}{N_{i,1,1}} \right. \\
& \left. + \sum_{k=2}^Y \left\{ - (a_{i,j,k,1} + \rho_{i,k,1} a_{i,j,k,2}) \prod_{m=1}^{k-1} \prod_{b=1}^2 \rho_{i,m,b} + \frac{y_{i,j,k}}{N_{i,1,1}} \right\} \right],
\end{aligned}$$

and equating to zero we find

$$N_{i,1,1} = \sum_{j=1}^T \frac{\sum_{k=1}^Y y_{i,j,k}}{a_{i,j,1,1} + \rho_{i,j,1} a_{i,j,1,2} + \sum_{k=2}^Y \left\{ (a_{i,j,k,1} + \rho_{i,k,1} a_{i,j,k,2}) \prod_{m=1}^{k-1} \prod_{b=1}^2 \rho_{i,m,b} \right\}}. \quad (14)$$

We note again how  $N_{i,1,1}$  is a weighted sum over visits of totals at site  $i$  across years. As in the univoltine case, we substitute the expressions for  $\{N_{i,1,1}\}$  from (14) into (13) and maximise the overall concentrated likelihood with respect to parameters associated with  $\rho$  and  $\mathbf{a}$ . Estimation of  $\{N_{i,1,1}\}$  is obtained by substituting estimates of  $\{a_{i,j,k,b}\}$  and  $\{\rho_{i,k,b}\}$  into (14).

This concentrated likelihood approach applies for both the phenomenological and mechanistic models for bivoltine species, with variation only in the specification of  $\{a_{i,j,k,b}\}$ .

## REFERENCES

- Altermatt, F. (2010), "Climatic warming increases voltinism in european butterflies and moths", *Proceedings of the Royal Society B: Biological Sciences*, 277, 1281–1287.
- Baur, B., Coray, A., Minoretti, N. and Zschokke, S. (2005), "Dispersal of the endangered flightless beetle *Dorcadion fuliginator* (Coleoptera: Cerambycidae) in spatially realistic landscapes", *Biological Conservation*, 124, 49–61.
- Botham, M. S., Brereton, T. M., Middlebrook, I., Randle, Z. and Roy, D. B. (2013), United Kingdom Butterfly Monitoring Scheme report for 2012. Centre for Ecology & Hydrology, Wallingford.
- Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C. (1995), "A limited memory algorithm for bound constrained optimization", *SIAM Journal on Scientific Computing*, 16, 1190–1208.
- Calabrese, J. M. (2012), "How emergence and death assumptions affect count-based estimates of butterfly abundance and lifespan", *Population Ecology*, 54, 431–442.
- Dennis, E. B., Freeman, S. N., Brereton, T. and Roy, D. B. (2013), "Indexing butterfly abundance whilst accounting for missing counts and variability in seasonal pattern", *Methods in Ecology and Evolution*, 4, 637–645.
- Dennis, E. B., Morgan, B. J. T., Freeman, S. N., Brereton, T. and Roy, D. B. (2014), "A generalised abundance index for seasonal invertebrates", Technical report. University of Kent. <http://kar.kent.ac.uk/id/eprint/48967>.

- Dennis, R. L. H. and Sparks, T. H. (2007), "Climate signals are reflected in an 89 year series of British Lepidoptera records", *European Journal of Entomology*, 104, 763–767.
- Diamond, S. E., Cayton, H., Wepprich, T., Jenkins, C. N., Dunn, R. R., Haddad, N. M. and Ries, L. (2014), "Unexpected phenological responses of butterflies to the interaction of urbanization and geographic temperature", *Ecology*, 95, 2613–2621.
- Diamond, S. E., Frame, A. M., Martin, R. A. and Buckley, L. B. (2011), "Species' traits predict phenological responses to climate change in butterflies", *Ecology*, 92, 1005–1012.
- Freeman, S. N. and Newson, S. E. (2008), "On a log-linear approach to detecting ecological interactions in monitored populations", *Ibis*, 150, 250–258.
- Grill, A., Cerny, A. and Fiedler, K. (2013), "Hot summers, long life: egg laying strategies of *Maniola* butterflies are affected by geographic provenance rather than adult diet", *Contributions to Zoology*, 82, 27–36.
- Hindle, B. J., Kerr, C. L., Richards, S. A. and Willis, S. G. (2015), "Topographical variation reduces phenological mismatch between a butterfly and its nectar source", *Journal of Insect Conservation*, 19, 227–236.
- Hodgson, J. A., Thomas, C. D., Oliver, T. H., Anderson, B. J., Brereton, T. M. and Crone, E. E. (2011), "Predicting insect phenology across space and time", *Global Change Biology*, 17, 1289–1300.
- Isaac, N. J. B., Girardello, M., Brereton, T. M. and Roy, D. B. (2011), "Butterfly abundance in a warming climate: patterns in space and time are not congruent", *Journal of Insect Conservation*, 15, 233–240.
- Johnson, D. S., Conn, P., Hooten, M., Ray, J. and Pond, B. A. (2012), "Spatial occupancy models for large data sets", *Ecology*, 94, 801–808.
- Karlsson, B. (2014), "Extended season for northern butterflies", *International Journal of Biometeorology*, 58, 691–701.
- Lahoz-Monfort, J. J., Morgan, B. J. T., Harris, M. P., Daunt, F., Wanless, S. and Freeman, S. N. (2013), "Breeding together: modeling synchrony in productivity in a seabird community", *Ecology*, 94, 3–10.
- Lahoz-Monfort, J. J., Morgan, B. J. T., Harris, M. P., Wanless, S. and Freeman, S. N. (2011), "A capture–recapture model for exploring multi-species synchrony in survival", *Methods in Ecology and Evolution*, 2, 116–124.
- Matechou, E., Dennis, E. B., Freeman, S. N. and Brereton, T. (2014), "Monitoring abundance and phenology in (multivoltine) butterfly species: a novel mixture model", *Journal of Applied Ecology*, 51, 766–775.
- Nowicki, P., Bonelli, S., Barbero, F. and Balletto, E. (2009), "Relative importance of density-dependent regulation and environmental stochasticity for butterfly population dynamics", *Oecologia*, 161, 227–239.
- Pagel, J., Anderson, B. J., O'Hara, R. B., Cramer, W., Fox, R., Jeltsch, F., Roy, D. B., Thomas, C. D. and Schurr, F. M. (2014), "Quantifying range-wide variation in population trends from local abundance surveys and widespread opportunistic occurrence records", *Methods in Ecology and Evolution*, 5, 751–760.
- Parker, D. E., Legg, T. P. and Folland, C. K. (1992), "A new daily Central England Temperature series, 1772–1991", *International Journal of Climatology*, 12, 317–342.
- Pereira, H. M., Leadley, P. W., Proença, V., Alkemade, R., Scharlemann, J. P., Fernandez-Manjarrés, J. F., Araújo, M. B., Balvanera, P., Biggs, R., Cheung, W. W. et al. (2010), "Scenarios for global biodiversity in the 21st century", *Science*, 330, 1496–1501.
- Powney, G. D., Roy, D. B., Chapman, D. and Oliver, T. H. (2010), "Synchrony of butterfly populations across species' geographic ranges", *Oikos*, 119, 1690–1696.
- R Core Team (2015), R: A language and environment for statistical computing. Vienna, Austria. <http://www.R-project.org/>.
- Rothery, P. and Roy, D. B. (2001), "Application of generalized additive models to butterfly transect count data", *Journal of Applied Statistics*, 28, 897–909.
- Roy, D. B., Rothery, P., Moss, D., Pollard, E. and Thomas, J. A. (2001), "Butterfly numbers and weather: predicting historical trends in abundance and the future effects of climate change", *Journal of Animal Ecology*, 70, 201–217.
- Roy, D. B. and Sparks, T. H. (2000), "Phenology of British butterflies and climate change", *Global Change Biology*, 6, 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, 832–841.
- Sparks, T. H. and Yates, T. J. (1997), "The effect of spring temperature on the appearance dates of British butterflies 1883–1993", *Ecography*, 20, 368–374.
- Spieth, H. R., Pörschmann, U. and Teiwes, C. (2011), "The occurrence of summer diapause in the large white butterfly *Pieris brassicae* (Lepidoptera: Pieridae): A geographical perspective", *European Journal of Entomology*, 108, 377–384.
- Thomas, C. D., Cameron, A., Green, R. E., Bakkenes, M., Beaumont, L. J., Collingham, Y. C., Erasmus, B. F., De Siqueira, M. F., Grainger, A., Hannah, L. et al. (2004), "Extinction risk from climate change", *Nature*, 427, 145–148.
- Thomas, J. A. (2005), "Monitoring change in the abundance and distribution of insects using butterflies and other indicator groups", *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360, 339–357.
- Van Dyck, H., Bonte, D., Puls, R., Gotthard, K. and Maes, D. (2015), "The lost generation hypothesis: could climate change drive ectotherms into a developmental trap?", *Oikos*, 124, 54–61.