




A Generic Method for Estimating and Smoothing Multispecies Biodiversity Indicators Using Intermittent Data

Stephen N. FREEMAN, Nicholas J. B. ISAAC , Panagiotis BESBEAS, Emily B. DENNIS, and Byron J. T. MORGAN

Biodiversity indicators summarise extensive, complex ecological data sets and are important in influencing government policy. Component data consist of time-varying indices for each of a number of different species. However, current biodiversity indicators suffer from multiple statistical shortcomings. We describe a state-space formulation for new multispecies biodiversity indicators, based on rates of change in the abundance or occupancy probability of the contributing individual species. The formulation is flexible and applicable to different taxa. It possesses several advantages, including the ability to accommodate the sporadic unavailability of data, incorporate variation in the estimation precision of the individual species' indices when appropriate, and allow the direct incorporation of smoothing over time. Furthermore, model fitting is straightforward in Bayesian and classical implementations, the latter adopting either efficient Hidden Markov modelling or the Kalman filter. Conveniently, the same algorithms can be adopted for cases based on abundance or occupancy data—only the subsequent interpretation differs. The procedure removes the need for bootstrapping which can be prohibitive. We recommend which of two alternatives to use when taxa are fully or partially sampled. The performance of the new approach is demonstrated on simulated data, and through application to three diverse national UK data sets on butterflies, bats and dragonflies. We see that uncritical incorporation of index standard errors should be avoided.

Supplementary materials accompanying this paper appear online.

Key Words: Bats; Butterflies; Dragonflies; Hidden Markov models; Hierarchical models; State-space models.

S. N. Freeman · N. J. B. Isaac (✉), UK Centre for Ecology & Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire, UK (E-mail: njbi@ceh.ac.uk). P. Besbeas, Department of Statistics, Athens University of Economics and Business, Athens, Greece. P. Besbeas · E. B. Dennis · B. J. T. Morgan, School of Mathematics, Statistics and Actuarial Science, University of Kent, Canterbury, Kent, UK. E. B. Dennis, Butterfly Conservation, Manor Yard, East Lulworth, Wareham, Dorset, UK.

© 2020 The Author(s)

Journal of Agricultural, Biological, and Environmental Statistics, Volume 26, Number 1, Pages 71–89
<https://doi.org/10.1007/s13253-020-00410-6>

1. INTRODUCTION

1.1. INDICATORS

Biodiversity indicators are one of the most direct ways in which biodiversity data inform governmental policy and action (Butchart et al. 2010; Costelloe et al. 2016), particularly in response to international targets of reducing the rate of biodiversity loss. Most multi-species indicators (MSIs) are based on small numbers of species (<50) within a narrowly defined scope (e.g. birds, butterflies or bats), although some combine species across a much wider taxonomic range (Eaton et al. 2015; van Strien et al. 2016). The transition to a new UN ‘Decade on Biodiversity’ is expected to stimulate the creation of new multispecies biodiversity indicators.

The UK has a long-established set of biodiversity indicators. The UK government reports annually on 11 MSIs, each spanning several decades. All report trend statistics which are in some sense ‘averages’ of time series for individual species. Most are based on count data and thus assumed to index changes in national abundance, although two are derived from occupancy detection models for several hundreds of species (Isaac et al. 2016). Generally, source data are gathered annually for each species at a large number of sampling sites and converted into time trends via statistical models containing ‘year’ effects that serve as indices of annual status. The precise way in which this is done varies between taxonomic groups for reasons of species ecology and sampling design. The methods of this paper do not require standardisation in this respect, and we shall not elaborate further on the many variations currently adopted for estimating individual species’ trends.

We note four points of inconsistency in contemporary indicators, both in the UK and internationally:

- Indicators are not consistent in the sources of uncertainty that are propagated through to estimates of precision in the headline MSI, and there are differences in the way in which precision is calculated.
- How, or indeed whether, the indicator line is smoothed.
- Indicators differ in how they deal with missing values, i.e. species:year combinations where there were insufficient surveys to produce reliable data.
- Whether the indicator measures changes in species abundance, species distributions (occupancy), or both.

1.2. DESIRABLE FEATURES OF MULTISPECIES INDICATORS

Buckland et al. (2005) discuss six useful properties a multispecies indicator should possess. We note two further specific qualities in addition to the issues raised above.

- Where practical, MSIs should follow a consistent methodology in order to ensure comparability across indicators and increase transparency. Harmonisation of indicator methods does not mean that we should also harmonise the methods used to calculate trends for individual species. Harmonising indicator methods but not the input data

implies a separation of indicator production from the models used to produce species indices. We can think of species index values derived from raw data as Essential Biodiversity Variables (Kissling et al. 2018), from which the indicator is a derived summary.

- The indicator should be interpretable in terms of changes in some meaningful biological unit, e.g. the ‘average’ change in abundance across species.

Differences across MSIs in statistical methods make it hard to compare across taxa. There is also a transparency deficit: MSIs are described in varying levels of detail, often in hard-to-find technical reports. These issues will become exacerbated as new indicators are developed that span multiple taxonomic groups with different time spans, methodologies and currencies (Eaton et al. 2015; van Strien et al. 2016).

1.3. PAPER OUTLINE

Our approach is based on a stochastic model for population growth rates, a convenient way to address the problem of missing species in particular years. In Sect. 2, we provide a hierarchical model for the calculation of multispecies biodiversity indicators within a state-space formulation. Classical and Bayesian implementations are given, based on the likelihood for MCMC and also the Kalman filter form, which is efficient for analysis using classical inference. We also provide smoothing within the indicator, rather than post hoc as is frequently done, with implications for the estimation of precision. Performance is demonstrated in the simulation study of Sect. 3. Section 4 provides a Bayesian application to British bat, butterfly and dragonfly data, and Sect. 5 illustrates the use of the classical approach for model selection. The paper ends with discussion, emphasising the wide applicability of the methods of the paper, and how they may be extended, in Sect. 6.

2. INFERENCE

2.1. MODEL FORMULATION

State-space models provide a convenient way to produce a multispecies indicator. Species index values can be expressed as a realisation of some unknown latent state, with a hierarchical structure that permits the full propagation of uncertainty. Sampling variances at the various levels are estimated, leaving trends that account for both sampling/observation error and population stochasticity.

Let $Y_{s,t}$ be the observed index of status for species s in year t , measured on the logarithmic scale, for $s = 1, \dots, n$ and $t = 1, \dots, T$. For example, Y could be the fitted log abundance derived from a Poisson linear model (Dennis et al. 2016b). However, working on the unbounded scale means that Y could equally represent the log odds of a site being occupied, derived from a probability-based site occupancy model (Dennis et al. 2017; MacKenzie et al. 2017). We assume these observations are normally distributed around a

true, unknown, value $S_{s,t}$:

$$Y_{s,t} \sim N(S_{s,t}, \sigma_{s,t}^2), \quad t = 1, 2, \dots, T. \quad (1)$$

Initially, we assume that $\sigma_{s,t}^2 = \sigma^2$, a constant to be estimated. In practice, this variation may change between species or years, and the model is readily extended in this fashion. Species-specific annual growth rates, $R_{s,t}$ follow from the relationships:

$$R_{s,t} = S_{s,t+1} - S_{s,t},$$

so that

$$S_{s,t} = S_{s,1} + \sum_{j=1}^{t-1} R_{s,j}, \quad t = 2, 3, \dots, T, \quad s = 1, \dots, n. \quad (2)$$

The indices in the first year could be regarded as unknown parameters or more parsimoniously taken as, for example

$$S_{s,1} \sim N(M_1, \tau_0^2), \quad s = 1, \dots, n, \quad (3)$$

with parameters M_1 and τ_0 to be estimated. Alternatively, species indices are often presented as, and always convertible to, trends relative to some value (e.g. 0, on this logarithmic scale) in the first year or, for some species, a subsequent year F_s . In this case, the option exists then to constrain $S_{s,1}$ or S_{s,F_s} also to zero. We assume the $\{R_{s,t}\}$ are normally distributed around multispecies average rates of growth, $\{G_t\}$, such that:

$$R_{s,t} \sim N(G_t, \tau_t^2), \quad t = 1, 2, \dots, T - 1, \quad s = 1, \dots, n. \quad (4)$$

This model recognises that growth rates vary among species each year, but that due to shared environmental conditions they may be generally higher or lower in one year than another. While their variance may also vary between years, throughout this paper we shall again assume for simplicity a constant value τ^2 .

Within this framework, we have two options for calculating a multispecies indicator. One metric, M_t , analogous to Eq. 2 can be derived by accumulating values of $\{G_t\}$:

$$M_t = M_1 + \sum_{j=1}^{t-1} G_j, \quad (5)$$

so that, with $M_1 = 0$, or estimated from the data, for example

$$G_t = M_{t+1} - M_t.$$

Where complete series are available for all species, common practice is to estimate the geometric mean abundance across them (Buckland et al. 2011), and following the same

logic, an alternative multispecies indicator may be derived as:

$$M'_t = \frac{1}{n} \sum_{s=1}^n S_{s,t}, \quad (6)$$

where n is the number of species. This is therefore the average of the individual species trends as estimated, on the log scale, within the model via Eq. 2. Presentation in the form $\exp(M'_t)$ gives geometric means of the fitted species indices, if these are similarly back-transformed. These model-derived species indices, $S_{s,t}$, will always exist, even for years in which an observed index, $Y_{s,t}$, is not available.

The unbounded scale, with Gaussian errors, is a convenient one in which to work and general enough to encompass either count or presence/absence data. Inverting the log transform produces an alternative presentation such as e^M , bounded only by zero, in which percentage change represents proportional change in abundance (count data) or the odds $\{p/(1-p)\}$ of species presence, consistent with standard practice, e.g. the Farmland Bird Indicator (Free-man et al. 2001). Further, indicators are often presented relative to zero (log-scale) or 100% in a base year, usually the first in the series, requiring just a final straightforward rescaling.

2.2. INFERENCE BASED ON THE KALMAN FILTER

The formulation of Eqs. 1–4 can alternatively be summarised by the equations:

$$S_{s,t+1} | S_{s,t} \sim \mathcal{N}(S_{s,t} + G_t, \tau_t^2) \quad (7)$$

$$S_{s,1} \sim \mathcal{N}(M_1, \tau_0^2) \quad (8)$$

$$Y_{s,t} | S_{s,t} \sim \mathcal{N}(S_{s,t}, \sigma_{s,t}^2). \quad (9)$$

The model can be seen to be a stochastic exponential growth model (Holmes 2001) and also a form of Gompertz state-space model on the log scale (Besbeas and Morgan 2020). Due to the normal distributions assumed, the model specified is a linear Gaussian state-space model. This is efficiently fitted using the Kalman filter, a recursive process with filtering and smoothing described in, for example, Newman et al. (2014, p.64), and we provide details in Sect. 2 of the Supplementary material. In the notation outlined there, the prediction error decomposition form of the log likelihood for a single species is given in Eq. 10. The likelihood is easily formed and maximum-likelihood estimation is fast. Simplicity of computation is clear from the short MATLAB program given in Besbeas and Morgan (2012). For species s , we denote the likelihood by L_s , and write

$$\log L_s(\boldsymbol{\theta}; Y_{s,1}, \dots, Y_{s,T}) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \left(\log |\mathbf{F}_t| + \mathbf{v}'_t \mathbf{F}_t^{-1} \mathbf{v}_t \right), \quad (10)$$

where $\boldsymbol{\theta}$ denotes the model parameters. The terms \mathbf{v}_t and \mathbf{F}_t are, respectively, the one-step forecast error of $Y_{s,t}$ given all earlier observations, and its variance. They are calculated through the implementation of the Kalman filter; see Equation 1 of the Supplementary material. The overall log likelihood is then formed by summing the terms in Eq. 10 over s .

A more general modelling framework is provided using hidden Markov modelling. Both approaches deal easily with missing data, for instance, for the Kalman filter this is explained in Harvey (1996, p.144). See also Sect. 2 of the Supplementary material.

2.3. BAYESIAN INFERENCE

For Bayesian analysis, the likelihood for \mathbf{Y} with the parameters θ is given formally by

$$L(\theta; \mathbf{Y}) = \prod_{s=1}^n \prod_{t=1}^T \phi(Y_{s,t}; S_{s,t}, \sigma^2),$$

where $\phi(\cdot)$ denotes the standard Normal probability density function, and the $\{S_{s,t}\}$ result from the recursion of Eq. 2. While we have specified the model in terms of normal distributions, other options might be considered.

The model of Eqs. 1–4 is readily specified in a Bayesian context in the BUGS language and has been implemented in the R package ‘BRCindicators’ (August et al. 2020). Prior distributions were set for the error standard deviations, and all M_t (or alternatively G_t), respectively, gamma distributions or uniform distributions with a lower bound at zero for the former, and normal distributions with mean zero and large variance for the latter. Results have not been found sensitive to these choices. Sampling error, if it cannot be assumed as known, is estimated and any values of $\{Y_{s,t}\}$ missed at random (at least with respect to the true rates of growth) are also automatically accounted for.

2.4. SMOOTHING

Conventionally, indices in single-species (Fewster et al. 2000) or compound (Freeman et al. 2001) form are generally smoothed for presentation. This is done to retain the signal of long-term changes and remove ephemeral influences, such as the effect of severe weather. Our approach makes the adoption of smoothing in situ straightforward. Many approaches to smoothing are available. The $\{G_t\}$, for example, are parameters to be estimated, and can be smoothed within the model. In a classical context, we use P-splines, with a linear function and, for most examples, 12 knots (Gimenez et al. 2006). Estimated variances of the MSI terms can be obtained using the delta method. In the paper, we demonstrate the behaviour of two alternatives, P-splines, as mentioned above, and the JAGS-based penalised spline formulation of Crainiceanu et al. (2005); growth is constrained such that, retaining the notation of the original paper:

$$G_t = m(t, \beta_1, \beta_2, \tau_\beta) \tag{11}$$

where the parameters of the function m , along with time t , represent those inducing the smoothing. (Note that we have omitted the additional variance term σ_{eps}^2 from the formula adopted by Crainiceanu et al. (2005), as we account for stochasticity via the other error terms in Eqs. 1–4). In the Bayesian approach, we need to select priors for these additional parameters, rather than M or G , and the number of knots that control the extent of smoothing

imposed. Here too we follow [Crainiceanu et al. \(2005\)](#) and embed the code they provide directly into our own. Smoothing on the growth rates, rather than the indicator provides a convenient way to accommodate missing data, which are merely imputed. The indicators M_t and M'_t are naturally smoothed as a consequence.

[Fewster et al. \(2000\)](#) further adopted a classical method of estimating both standard errors and the second derivatives of their smoothed trends, via bootstrapping. [Freeman et al. \(2001\)](#) used this method for multispecies indicators; for each species in turn sites were resampled, with replacement, to produce a replicate bootstrapped trend for each species. These were then combined into a multispecies indicator, and the entire process repeated multiple times to produce a number of bootstrapped indicators, each containing one (and only one, different) trend for each species. Years in which the bootstrapped 95% CI of the second derivatives excluded zero were taken to indicate a significant change in direction or a significantly increased or reduced rate of change—acceleration, as distinct from velocity ([Buckland et al. 2005](#); [Fewster et al. 2000](#)).

The smoothing of M_t directly, or alternatively growth rates G_t in Eq. 5, allows us to produce analogous estimates of second derivatives via a difference-equation approximation, e.g. using sixth differences where possible:

$$\frac{d^2 M_t}{dt^2} \approx \frac{1}{180} (2M_{t+3} - 27M_{t+2} + 270M_{t+1} - 490M_t + 270M_{t-1} - 27M_{t-2} + 2M_{t-3}), \quad (12)$$

and lower-order alternative difference approximations at the beginning and end of the series, where the calculation of Eq. 12 cannot be executed (see [Fewster et al. 2000](#)). In a Bayesian context, credible intervals immediately arise, and again by analogy we might take those years in which zero lies outside the 95% credible interval to indicate years in which the data suggest a notable increase/decrease or change in direction.

3. SIMULATION STUDY

To investigate the performance of the method, we simulated a set of data $\{Y_{s,t}\}$ for 8 species over an assumed 30-year period, consistent with Eqs. 1–4 and using a randomly chosen sequence $M_1 \dots M_{30}$ and specified variance parameters. For simplicity, we assume a time-invariant constant σ , both in generating data and in fitting the model, within which it is assumed unknown and estimated from the data. The model of Eqs. 1–4, implemented here in the Bayesian framework, provides an accurate reconstruction of the true indicator, M_t (Fig. 1a,b), and with complete data is virtually indistinguishable from the traditional geometric mean method.

However, it will be clear that if particular species indices are unavailable in certain years, bias will ensue if the absent data are not missing at random but, for example, predominantly rare (or rarely encountered) species, as may be the case in practice. Figure 1c shows a repeat of the analysis of Fig. 1a,b but with those data with the lowest values omitted (20% of the whole) and regarded as missing.

It is clear that, as most missed values are late in the series for this set of largely declining species, the geometric mean of the remaining data underestimates the strength of the decline,

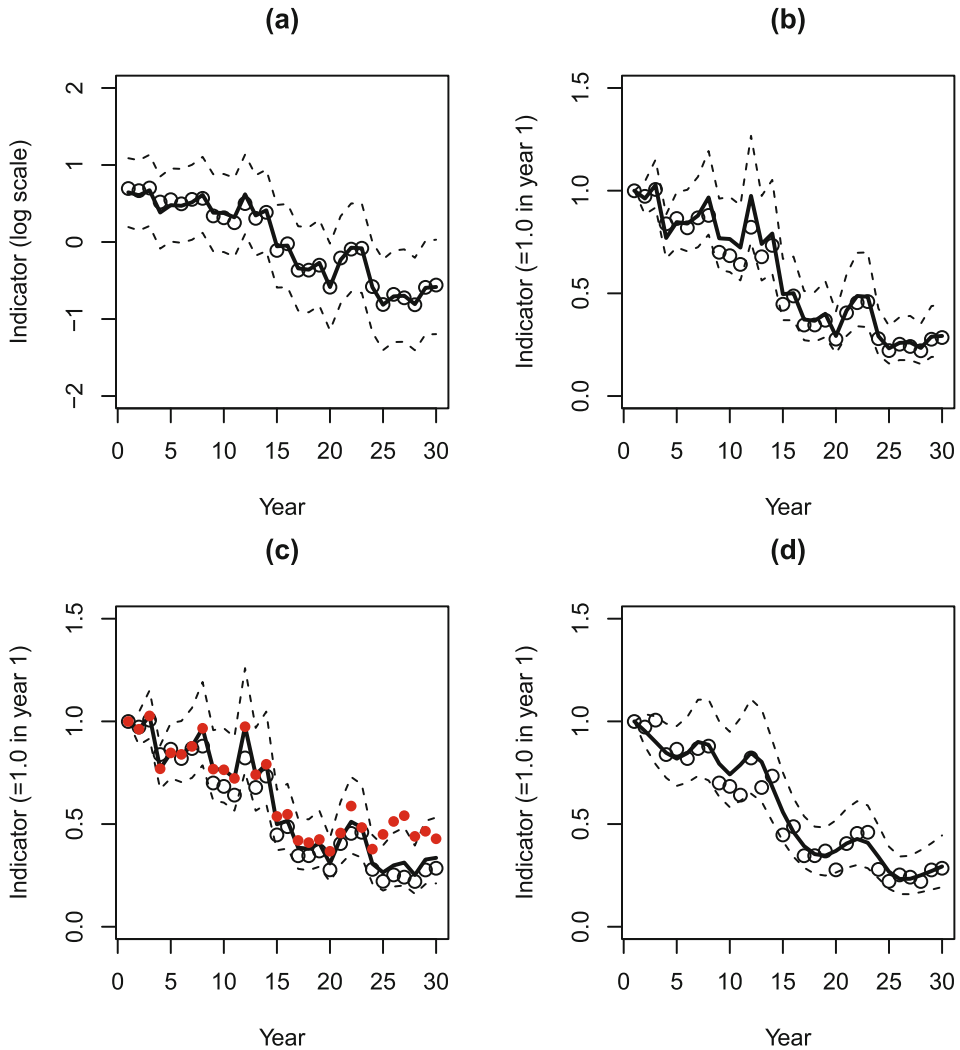


Figure 1. Simulation results — multispecies indicators: **a** M_t (solid line), based on all data and shown on original log scale, with 95% credible intervals; **b** $\exp(M_t)$, relative to baseline value of 1.0 in year 1 after transformation; **c** $\exp(M_t)$ with lowest 20% of data omitted geometric mean of remaining data plotted in closed circles; **d** Smoothed version of $\exp(M_t)$, all data. True values used in simulation are plotted as open circles throughout.

while the growth-based alternative remains robust. Parameter values used in the simulations and a summary from a larger number of simulations are shown in Supplementary Figures 1 and 2.

We then applied the smoother of Eq. 11 to the entire dataset, with missing values reinstated. Priors for smoothing parameters were exactly as in Crainiceanu et al. (2005), and twelve knots were used throughout. The smoother accurately captures the changes in the indicator, and the species-specific components, whose growth rates are assumed normally distributed about $\{G_t\}$ are also estimated accurately (Figs. 1d, 2).

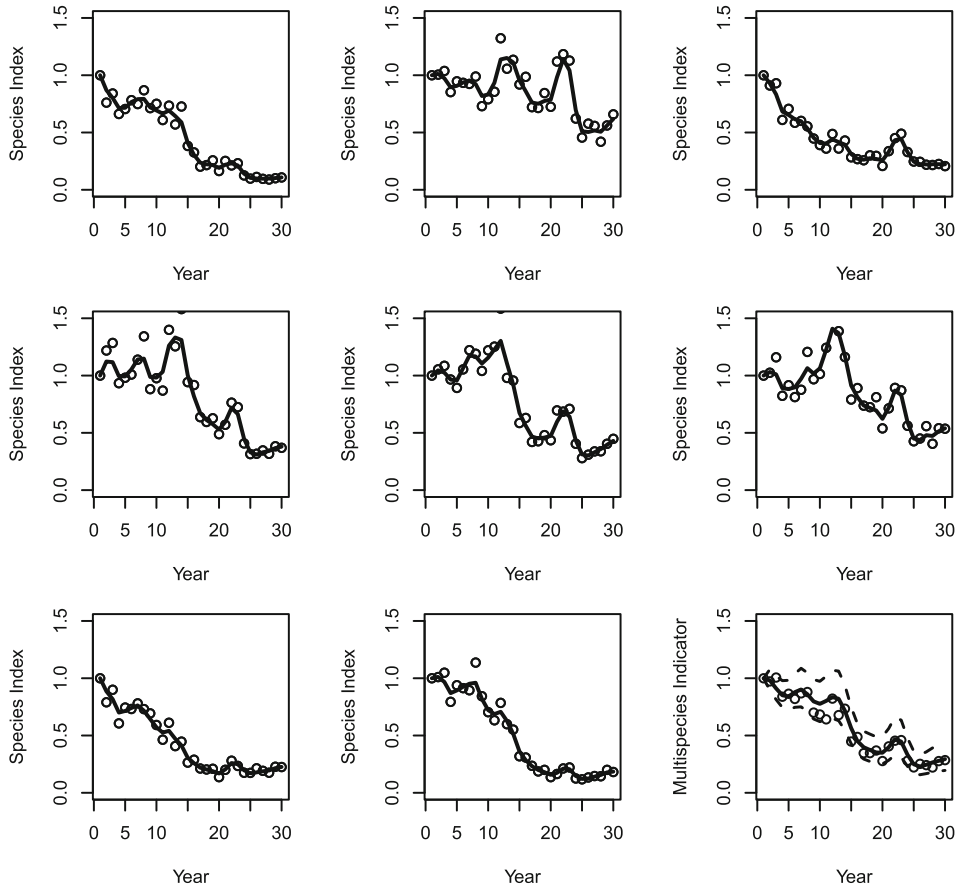


Figure 2. Simulation results. True indices of abundance for eight species (open circles) and fitted values (solid line). For ease of comparison Figure 1d is repeated in the bottom right-hand corner.

4. APPLICATION: UK MULTISPECIES INDICATORS

We present four case studies to illustrate the flexibility of our method. In each case, we present results from the Bayesian implementation, but the results from the Kalman filter were indistinguishable.

4.1. UK BATS

There are 17 species of bat breeding in the UK. Annual indices for eight of these species, produced by the National Bat Monitoring Program, contribute to UK biodiversity indicators (Supplementary Table 1). The rest have limited data or are, in the case of several *Myotis* species, difficult to identify to species level. The data are based on counts from a variety of surveys, including direct observations at hibernation roosts and field surveys using bat detectors, which are combined into a single index for each species. Six of the eight species

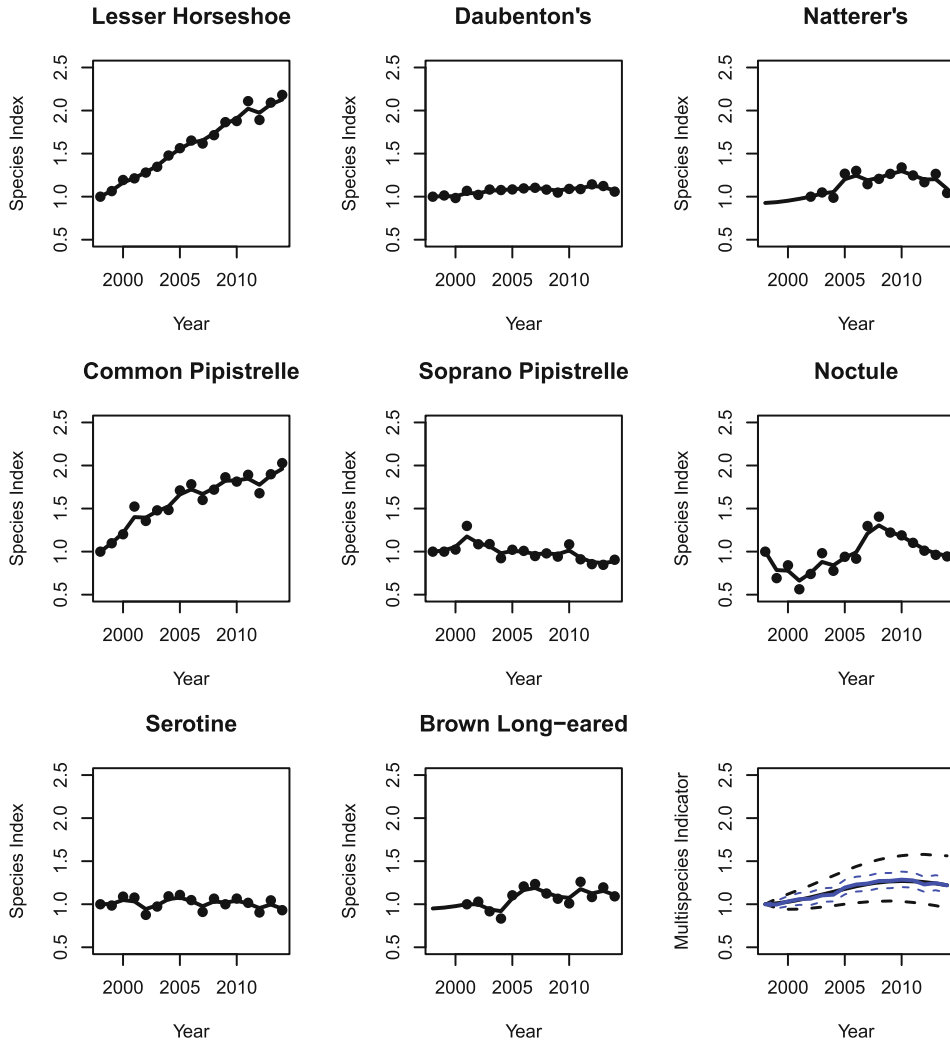


Figure 3. Observed (closed circles) and fitted (solid line) species indices for 8 British bat species. Corresponding (smoothed) multispecies indicators based on M_t and M'_t are shown in bottom right-hand corner — indicator $\exp(M_t)$ is shown in black, geometric means $\exp(M'_t)$ in grey (blue in online version), both with 95% credible intervals and each back-transformed and scaled to 1.0 in 1998, or individual species' year of first encounter.

have index values for every year (1998–2014): the other two enter the time series in 2001 (Brown Long-eared) and 2002 (Natterer's).

We fitted the model of Eqs. 1–4 to these data using JAGS. The indicator shows an increase over the duration of this relatively short series, and changes smoothly, with no significant second derivatives, as a consequence of the limited ability for rapid increase in species with low annual productivity (Fig. 3).

4.2. UK BUTTERFLIES

There are 59 regularly occurring species of butterfly in the UK, and data from the UK Butterfly Monitoring Scheme (UKBMS) produces annual indices for 57 of these (Brereton et al. 2018). The data are primarily based on counts collected from standardised transects, as well as reduced effort methods including timed counts, larval web counts and egg counts. Since 2009, additional data have been contributed from the Wider Countryside Butterfly Survey (Brereton et al. 2011). Species-specific indices are derived using the approach of Dennis et al. (2016a), with an additional modification—see Brereton et al. (2018).

Indices from UKBMS data are typically separated into habitat specialists (26 species) and wider countryside species (24 species). The latter has only one species (Scotch Argus) that joins the series late, whereas the habitat specialists dataset is far more incomplete, with only ten species possessing an index value in every year. Another 15 species were added to the series between 1978 and 1995, and one species (Swallowtail) has data for every year except 1978. Full species lists for both datasets are listed in Supplementary Table 2.

We used the model of Eqs. 1–4 to produce both Bayesian smoothed indicators for both datasets; see Fig. 4. The fitted species indices (Supplementary Figures 3 and 4) demonstrate how the model smooths the population fluctuations of individual species' time series.

For both butterfly models in Fig. 4, years are indicated in which the second derivative of the growth rates are markedly nonzero. Clearly, the trend for wider countryside species is considerably more variable over short-term periods, even after smoothing, although both have stabilised over the last decade. Over the duration of the series however, the decrease has been greater for the specialist species. For both datasets, there is little difference between M_t and M'_t , although the geometric mean method (M'_t) is very much more precise.

4.3. UK DRAGONFLIES

Dragonfly distribution records have been collected by the Dragonfly Recording Network and coordinated by the British Dragonfly Society. Dragonflies are among the best-recorded insect taxa in the UK, after butterflies (Isaac and Pocock 2015; Powney et al. 2015). Although these data constitute presence-only records, species trends can be estimated using occupancy-detection models (MacKenzie et al. 2017). We took advantage of existing occupancy-detection models for UK dragonflies (Outhwaite et al. 2019), which had been fitted using methods described in Isaac et al. (2014) and Outhwaite et al. (2018). From these, we extracted occupancy estimates for each year 1970–2015, on the unbounded logit scale before calculation of the Bayesian MSI, which are shown in Supplementary Figure 5. For the multispecies indicator, we use the 38 species (listed in Supplementary Table 3) for which trends were estimated in the most recent Atlas of Dragonflies (Cham et al. 2014), with the exception of Irish Damselfly, which does not occur in Great Britain. The time series are complete except Small red-eyed Damselfly, for which the first estimate is in 2001, the year in which it was first recorded in Great Britain.

We fitted the smoothed multispecies model (Fig. 5, with fitted species trends in Supplementary Figure 5). A general increase in the multispecies indicators from the late 1980s is revealed, with several spells of rapid average expansion, but with a reversal of direction

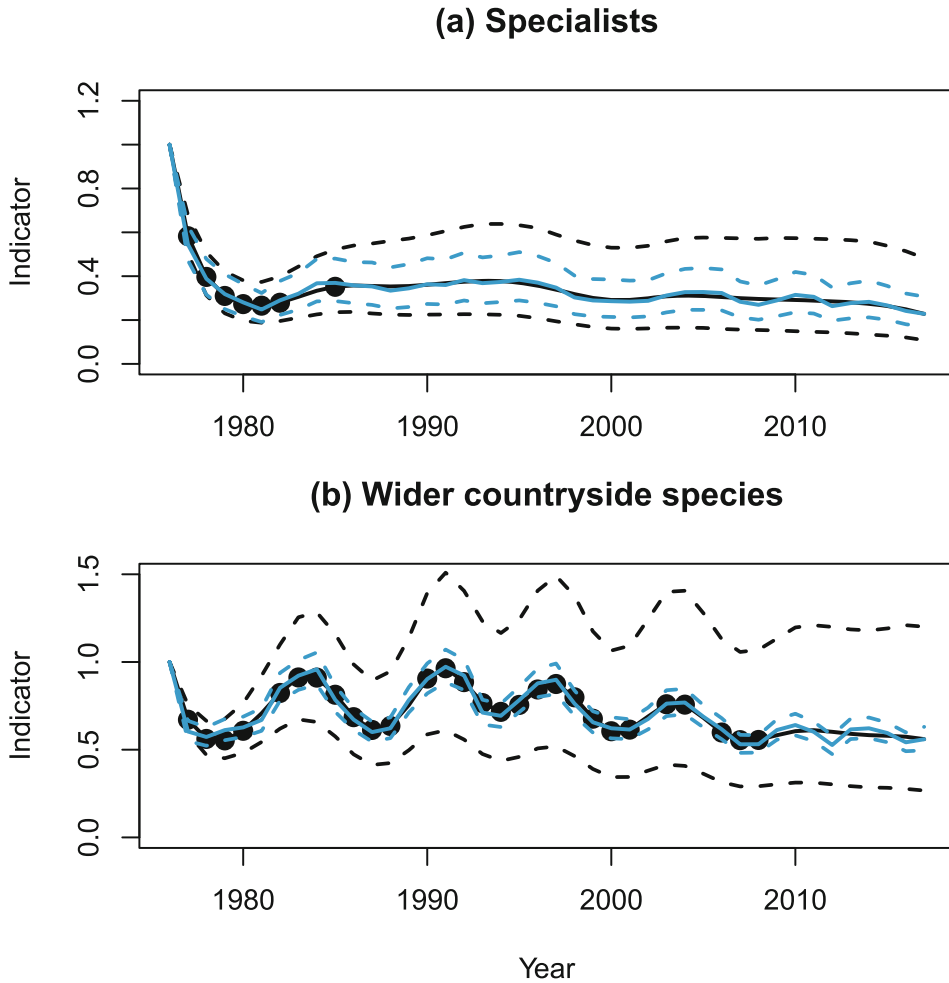


Figure 4. Smoothed butterfly indicators, **a** habitat specialist species and **b** Wider countryside species. Indicator $\exp(M'_t)$ is shown in black, geometric means $\exp(M'_t)$ in grey (blue in online version), both with 95% credible intervals and each transformed to 1.0 in 1976. Filled circles indicate years in which the credible intervals of the second derivative of M'_t do not include zero.

towards the very end of the series. Note that the indicators, based on changes in range, are smooth and very much more so, with a similar number of knots, than those in the abundance-based wider countryside species butterfly graph, though more similar to that for the habitat specialists. Fitted species trends, however, are noticeably less smooth than their counterparts in the analogous butterfly models.

5. INCLUDING ESTIMATED STANDARD ERRORS

Individual species indices may arise with associated measures of uncertainty, for example in the form of standard errors $\{se_{s,t}\}$, which may be obtained in different ways. These have been regarded as difficult to incorporate directly (Soldaat et al. 2017) but, in obvious notation,

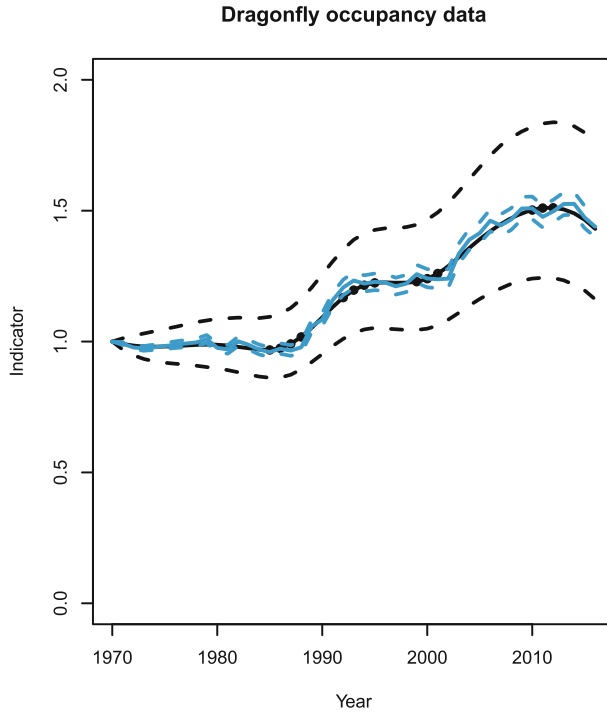


Figure 5. Dragonfly indicator, based upon occupancy data. Lines and symbols as Fig. 4.

the proposed methodology allows us to model observation error variance $\{\sigma_{s,t}^2\}$ in terms of estimated standard error using various formulations, along a spectrum ranging from values fixed a priori to the estimation of entirely free parameters, e.g.:

$$\begin{aligned}
 \text{[i]} \quad & \sigma_{s,t}^2 = 0 \\
 \text{[ii]} \quad & \sigma_{s,t}^2 = se_{s,t}^2 \\
 \text{[iii]} \quad & \sigma_{s,t}^2 = c^2 se_{s,t}^2 \\
 \text{[iv]} \quad & \sigma_{s,t}^2 = \theta^2,
 \end{aligned}$$

where θ and c are unknown non-negative parameters to be estimated. Model iv, the most general, is our basic model of Eqs. 1–4.

We can compare these alternative models relying solely on information in the data using classical inference (i.e. the Kalman filter). For example, in Table 1 we compare the above four models fitted to the UKBMS data, and the dragonfly data and models (i) and (iv) for the bat data. UKBMS and dragonfly data were smoothed via a P-spline with 12 knots, and bats with 5 knots (reflecting a shorter time series for this taxon). It is interesting that the best model for the UKBMS data is the model with a constant observation variance, suggesting that the estimated standard errors are in conflict with the model and the indices in this case, despite the use of a scaling factor to account for standardisation. This may be because the estimates of error at the species level are derived from one stage of a statistical model and are therefore subject to additional uncertainty that we have not accounted for; see Dennis

Table 1. A comparison of the four alternative models for the observation error applied to the full set of UKBMS data, for the two subsets: wider countryside and habitat specialists, the dragonfly data and also the bat data

Model	Butterflies			Dragonflies	Bats
	max logL			max logL	max logL
	Full	Wider	Specialists		
(i)	1949.2	1024.6	939.7	2974.6	208.3
(ii)	1999.3	1074.8	937.8	2536.2	–
(iii)	2004.8	1079.3	958.7	2977.4	–
(iv)	2114.7	1117.8	1022.9	2974.6	222.5

We denote overall likelihood by L . Model (i): $\sigma_{s,t}^2 = 0$; model (ii): $\sigma_{s,t}^2 = se_{s,t}^2$; model (iii): $\sigma_{s,t}^2 = c^2 se_{s,t}^2$; model (iv): $\sigma_{s,t}^2 = \theta^2$. In each column, the largest value of the maximum log likelihood values is shown in bold. In this illustration, the Kalman filter has been used

et al. (2016a) for details. The results for bats are limited, as estimates of standard error at the species level are not available. For the dragonfly data, models (i) and (iv) give the same parameter estimates, as under model (iv) θ is estimated on the boundary $\theta = 0$. This is because there is little variation between the indices in this application, and presumably whatever variation there is is being taken up through the system part of the model, rather than the observation process. Plugging in the estimated index standard errors directly, as in model (ii), therefore results in a large drop in maximised log likelihood. The best model is model (iii) as the scaling constant c can greatly reduce the estimated standard errors, while maintaining variation in the estimation precision of the indices. An interesting feature of the dragonfly data is that one species, Small Red-eyed Damselfly, was a late entrant, with data first recorded in 2001. However, the indices that we used were the result of the model of Outhwaite et al. (2019), which estimates missing data. If we use the model of this paper so that for that species we only use records from 2001, then there is no difference between models (iii) and (iv), as in each case $\sigma_{s,t}$ is estimated on the boundary $\sigma_{s,t} = 0$.

6. DISCUSSION

6.1. ADVANTAGES OF THE NEW MODEL

Use of a state-space model for an MSI meets the prevailing need for transparent, consistent and flexible summaries of data. It is based upon reasonable assumptions about variation in annual change between species. A key advantage is that it allows a degree of standardisation in the production of multispecies indicators, while permitting flexibility in the derivation of the single-species trends contributing to them. Notably, it is applicable either to trends derived from count data or to trends in occupancy. Note that combining trends in occupancy and abundance into a *single* MSI may not be appropriate, as changes in species' abundance do not generally equate directly to changes in distribution size (Dennis et al. 2019).

The stochastic, state-space form allows separation of sampling errors and population stochasticity, with flexibility in the way that both are handled. The method is robust to

values missed not at random and, specifically, it readily accommodates species for which appropriate data arrive only late in the series, either because these are recent, rapid colonists like the Small Red-eyed Damselfly or because data from the early period is considered inadequate (as in the UKBMS). The model improves upon the use of the geometric mean when data are missing, as seen in Fig. 1. Applied to three very different datasets, the model provides both a good fit to the observed species' indices and an all-species indicator smooth enough to capture major changes and identify points of transition.

Smoothing in situ is a particular advantage where the available species indices are not smoothed (as in the UK butterflies) or where they contain a mixture of smoothed and unsmoothed data (as in the UK Priority Species Indicator, Eaton et al. 2015). Credible/confidence limits are immediately tractable, with or without smoothing, without the often prohibitive need for bootstrapping. Such limits for derivatives of the smoothed trend are equally straightforward, at least in a Bayesian context. Note that smoothing requires the user to define the number of knots, which is effectively a free parameter. We followed Crainiceanu et al. (2005) and Fewster et al. (2000) in selecting the number of knots in our examples, but the resulting MSIs were very similar unless the number of knots was changed radically (results not shown).

The approach also permits the direct incorporation of existing estimates of observation error. However, our results suggest that while this can be an improvement over disregarding such errors altogether, it is not always optimal. An important finding is that plugging in estimated standard errors should be carefully evaluated, as shown in Table 1.

Our approach is made possible by assuming that species' growth rates are normally distributed around some global average (Eq. 4). It is possible to envisage circumstances in which this assumption might be violated, leading to directional bias in the multispecies indicator. As with any complex model, we advocate post hoc diagnosis of the assumptions, e.g. by plotting the distribution of species-specific growth rates $\{R_{s,t}\}$. Violations of the normality assumption could be addressed by the addition of extra terms (e.g. via a mixture model).

Of the two alternative forms considered, that based on averages of the fitted species trends (M'_t) is more precise than that based on growth rates (M_t). Imprecision of (M'_t) reflects only the uncertainty in individual species' status, whereas (M_t) additionally reflects the fact that species in the dataset are assumed to be indicative of a wider set of species. The reality for most indicators is that species within are a highly non-random selection, included solely on the basis of data availability. This is exemplified by the case studies, which include the vast majority of bat, butterfly and dragonfly species in GB: the omissions are either rare, localised, cryptic or otherwise lacking in data. In these cases, M'_t seems an appropriate choice. The alternative M_t might be preferred where the species selected are a small fraction of the eligible set, and where extrapolation to the wider set of species is desirable. Illustrations based upon smaller subsets of data are given in Supplementary Figures 6-8.

6.2. FURTHER MODEL EXTENSIONS AND APPLICATIONS

Standardising to a baseline, standard practice among biodiversity indicators, may have undesirable consequences. First, it conveys the impression that precision is greatest near the

start of the series, typically the period with fewest data and therefore the lowest precision (e.g. the UKBMS contained just 30 sites in 1976, but >1000 in the last two decades). Second, it can lead to uncertainty growing over time; hence, counter-intuitively it becomes progressively harder to detect ‘significant’ change even as data increase (Gregory et al. 2019). Although we have followed this convention, minor modification would permit an indicator to be presented with any alternative baseline year, and/or uncertainty in the baseline year, at little additional computational cost.

We note, too, that the multispecies model of Eqs. 1–6 extends naturally to the initial estimation of growth rates and indices for a single species from data at multiple sites. Such data have format identical to that of the indices modelled in Eqs. 1–4, the subscript s merely indicates the site rather than species. If therefore $Y_{s,t}$ were redefined in such a situation as the count (say) of a particular species at site s in year t , application of the method produces M_t and M'_t that serve as indicators of average change in growth over the sites at which the species was observed. As the count in a raw data set will be a non-negative integer, often small or zero in practice, it is rational to amend Eq. 1 such that either

$$\log(1 + Y_{s,t}) \sim N(S_{s,t}, \sigma_{s,t}^2),$$

or

$$Y_{s,t} \sim \text{Poisson}\{\exp(S_{s,t})\}.$$

As the Y now represent absolute counts (rather than relative indices) the method would proceed without any standardisation to an arbitrary value in a base year, and the alternative indicator in Eq. 6 is perhaps more rationally replaced by the conventional arithmetic mean (or sum) of the fitted counts. In this context, the value is that the method does not make the assumption that the trend is the same at all sites, as frequently done to produce species indices used for UK indicators, or determined via a spatial covariate (Freeman and Newson 2008; Harrison et al. 2014; Dennis et al. 2016b). The model is thereby more flexible, and estimated trends for different sites can be meaningfully compared. These estimated species-specific M_t or M'_t could then be naturally combined into a multispecies indicator by treating them now as data and applying Eqs. 1–6. It remains, however, fundamental to the value of the approach that it is not necessary for all species indices combined into an indicator to have been initially calculated the same way, only that each is considered proportional, as appropriate, to abundance or odds of occupancy.

The model of the paper has been fitted using both MCMC and maximum-likelihood. Results show excellent agreement, providing a check of the methods. General advice on selecting a model-fitting procedure is provided by Newman et al. (2014, p. 80). The MCMC implementation has been incorporated into the R package ‘BRCindicators’ (August et al. 2020). Example code for both implementations is available as supplementary material.

ACKNOWLEDGEMENTS

SF and NI are funded by Natural Environment Research Council award NE/R016429/1 as part of the UK-SCAPE programme delivering National Capability. PB was partly supported by an Original Research Grant, AUEB. BJTM

was supported by a Leverhulme research fellowship. The UKBMS is run by Butterfly Conservation, the UK Centre for Ecology & Hydrology, and the British Trust for Ornithology, in partnership with a consortium of government agencies. We are grateful to Tom August for technical support, and to all the volunteers who have gathered data for the multiple taxa featured in this paper. We thank Katherine Boughey and two anonymous reviewers for comments on a previous version of this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

[Received August 2019. Accepted July 2020. Published Online August 2020.]

REFERENCES

- August, T. A., Powney, G., Outhwaite, C. L., Hatfield, J., Logie, M., Freeman, S. N., et al. (2020). BRCindicators: creating multispecies biodiversity indicators. R package version 1.3.6. <https://github.com/BiologicalRecordsCentre/BRCindicators>.
- Barlow, K. E., Briggs, P. A., Haysom, K. A., Hutson, A. M., Lechiara, N. L., Racey, P. A., et al. (2015). Citizen science reveals trends in bat populations: the National Bat Monitoring Programme in Great Britain. *Biological Conservation*, 182:14–26.
- Besbeas, P. and Morgan, B. J. T. (2012). Kalman filter initialization for integrated population modelling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 61(1):151–162.
- Besbeas, P. T. and Morgan, B. J. T. (2020). A general framework for modelling population abundance data. *Biometrics*, 76:281–292. <https://doi.org/10.1111/biom.13120>.
- Brereton, T. M., Botham, M. S., Middlebrook, I., Randle, Z., Noble, D., Harris, S., et al. (2018). United Kingdom Butterfly Monitoring Scheme report for 2017. Technical report, Centre for Ecology & Hydrology, Butterfly Conservation, British Trust for Ornithology and Joint Nature Conservation Committee.
- Brereton, T. M., Cruickshanks, K. L., Risely, K., Noble, D. G., and Roy, D. B. (2011). Developing and launching a wider countryside butterfly survey across the United Kingdom. *Journal of Insect Conservation*, 15:279–290.
- Buckland, S. T., Magurran, A. E., Green, R. E., and Fewster, R. M. (2005). Monitoring change in biodiversity through composite indices. *Phil. Trans. R. Soc. B*, 360:243–254.
- Buckland, S. T., Studeny, A. C., Magurran, A., Illian, J. B., and Newson, S. (2011). The geometric mean of relative abundance indices: a biodiversity measure with a difference. *Ecosphere*, 2. <https://doi.org/10.1890/ES11-00186.1>.
- Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R., et al. (2010). Global biodiversity: indicators of recent declines. *Science*, 328(5982):1164–1168. <https://doi.org/10.1126/science.1187512>.
- Cham, S. A., Nelson, B., Parr, A., Prentice, S., Smallshire, D., and Taylor, P. (2014). *Atlas of Dragonflies in Britain and Ireland*. Field Studies Council.
- Costelloe, B., Collen, B., Milner-Gulland, E. J., Craigie, I. D., McRae, L., Rondinini, C., and Nicholson, E. (2016). Global biodiversity indicators reflect the modeled impacts of protected area policy change. *Conservation letters*, 9(1):14–20.
- Crainiceanu, C., Ruppert, D., and Wand, M. P. (2005). Bayesian analysis for penalized spline regression using WinBUGS. *Journal of Statistical Software*, 14(14).

- Dennis, E. B., Brereton, T., Morgan, B. J., Fox, R., Shortall, C. R., Prescott, T., and Foster, S. (2019). Trends and indicators for quantifying moth abundance and occupancy in Scotland. *Journal of Insect Conservation*, 23(2):369–380.
- Dennis, E. B., Morgan, B. J. T., Freeman, S. N., Brereton, T., and Roy, D. B. (2016a). A generalized abundance index for seasonal invertebrates. *Biometrics*, 72(4):1305–1314.
- Dennis, E. B., Morgan, B. J. T., Freeman, S. N., Ridout, M. S., Brereton, T. M., Fox, R., et al. (2017). Efficient occupancy model-fitting for extensive citizen-science data. *PLoS ONE*. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0174433>.
- Dennis, E. B., Morgan, B. J. T., Freeman, S. N., Roy, D. B., and Brereton, T. M. (2016b). Dynamic models for longitudinal butterfly data. *Journal of Agricultural, Biological, and Environmental Statistics*, 21:1–21.
- Eaton, M. A., Burns, F., Isaac, N. J. B., Gregory, R. D., August, T. A., Barlow, K. E., et al. (2015). The priority species indicator: measuring the trends in threatened species in the UK. *Biodiversity*, 16(2-3):108–119.
- Fewster, R. M., Buckland, S. T., Siriwardena, G. M., Baillie, S. R., and Wilson, J. D. (2000). Analysis of population trends for farmland birds using generalized additive models. *Ecology*, 81(7):1970–1984.
- Freeman, S. N., Baillie, S. R., and Gregory, R. D. (2001). Statistical analysis of an indicator of population trends in farmland birds. *BTO Research Report*.
- Freeman, S. N. and Newson, S. E. (2008). On a log-linear approach to detecting ecological interactions in monitored populations. *Ibis*, 150:250–258.
- Gimenez, O., Crainiceanu, C., Barbraud, C., Jenouvrier, S., and Morgan, B. J. T. (2006). Semiparametric regression in capture-recapture modelling. *Biometrics*, 62:691–698.
- Gregory, R. D., Skorpilova, J., Vorisek, P., and Butler, S. (2019). An analysis of trends, uncertainty and species selection shows contrasting trends of widespread forest and farmland birds in Europe. *Ecological Indicators*, 103:676–687.
- Harrison, P. J., Buckland, S. T., Yuan, Y., Elston, D. A., Brewer, M. J., Johnston, A., and Pearce-Higgins, J. W. (2014). Assessing trends in biodiversity over space and time using the example of British breeding birds. *Journal of Applied Ecology*, 51(6):1650–1660.
- Harvey, A. C. (1996). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Cambridge.
- Holmes, E. E. (2001). Estimating risks in declining populations with poor data. *Proceedings of the National Academy of Sciences*, 98:5072–5077.
- Isaac, N., Eaton, M., Burns, F., Rob, D., Gregory, R., Peyton, J., and Roy, D. (2016). Provision of evidence statements to accompany the UK and England species trend indicators and an overview of the causes of biodiversity change. Final Report. Defra Project reference BE0112.
- Isaac, N. J. B. and Pocock, M. J. O. (2015). Bias and information in biological records. *Biological Journal of the Linnean Society*, 115(3):522–531.
- Isaac, N. J. B., van Strien, A. J., August, T. A., de Zeeuw, M. P., and Roy, D. B. (2014). Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution*, 5(10):1052–1060.
- Kissling, W. D., Ahumada, J. A., Bowser, A., Fernandez, M., Fernández, N., García, E. A., et al. (2018). Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. *Biological reviews*, 93(1):600–625.
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., and Hines, J. E. (2017). *Occupancy Estimation and Modeling, 2nd edition: Inferring Patterns and Dynamics of Species Occurrence*. Academic Press, New York.
- Newman, K. B., Buckland, S. T., Morgan, B. J. T., King, R., Borchers, D. L., Cole, D. J., et al. (2014). *Modelling Population Dynamics*. Springer, New York.
- Outhwaite, C. L., Chandler, R. E., Powney, G. D., Collen, B., Gregory, R. D., and Isaac, N. B. (2018). Prior specification in Bayesian occupancy modelling improves analysis of species occurrence data. *Ecological Indicators*, 93:333–343.

- Outhwaite, C. L., Powney, G. D., August, T. A., Chandler, R. E., Rorke, E., Pescott, O., et al. (2019). Annual estimates of occupancy for bryophytes, lichens and invertebrates in the UK (1970–2015). NERC Environmental Information Data Centre.
- Powney, G. D., Cham, S. S. A., Smallshire, D., and Isaac, N. J. B. (2015). Trait correlates of distribution trends in the Odonata of Britain and Ireland. *PeerJ*, 3:e1410.
- Soldaat, L. L., Pannekoek, J., Verweij, R. J. T., van Turnhout, C. A. M., and van Strien, A. J. (2017). A Monte Carlo method to account for sampling error in multi-species indicators. *Ecological Indicators*, 81:340–347. <https://doi.org/10.1016/j.ecolind.2017.05.033>.
- van Strien, A. J., Meyling, A. W. G., Herder, J. E., Hollander, H., Kalkman, V. J., Poot, M. J. M., et al. (2016). Modest recovery of biodiversity in a western European country: The Living Planet Index for the Netherlands. *Biological Conservation*, 200:44–50.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.