Journal of Applied Ecology

Journal of Applied Ecology 2016, **53,** 1852–1861 doi: 10.1111/1365-2664.12720

Decomposing trends in Swedish bird populations using generalized additive mixed models

Jonas Knape*

Department of Ecology, Swedish University of Agricultural Sciences, Box 7044, 750 07 Uppsala, Sweden

Summary

1. Estimating trends of populations distributed across wide areas is important for conservation and management of animals. Surveys in the form of annually repeated counts across a number of sites are used in many monitoring programmes, and from these, nonlinear trends may be estimated using generalized additive models (GAM).

2. I use generalized additive mixed models (GAMM) to decompose population change into a long-term, smooth, trend component and a component for short-term fluctuations. The long-term population trend is modelled as a smooth function of time and short-term fluctuations as temporal random effects. The methods are applied to analyse trends in goldcrest and greenfinch populations in Sweden using data from the Swedish Breeding Bird Survey. I use simulations to investigate statistical properties of the model.

3. The model separates short-term fluctuations from longer term population change. Depending on the amount of noise in the population fluctuations, estimated long-term trends can differ markedly from estimates based on standard GAMs. For the goldcrest with wide among-year fluctuations, trends estimated with GAMs suggest that the population has in recent years recovered from a decline. When filtering out, short-term fluctuations analyses suggest that the population has been in steady decline since the beginning of the survey.

4. Simulations suggest that trend estimation using the GAMM model reduces spurious detection of long-term population change found with estimates from a GAM model, but gives similar mean square errors. The simulations therefore suggest that the GAMM model, which decomposes population change, estimates uncertainty of long-term trends more accurately at little cost in detecting them.

5. *Policy implications*. Filtering out short-term fluctuations in the estimation of long-term smooth trends using temporal random effects in a generalized additive mixed model provides more robust inference about the long-term trends compared to when such random effects are not used. This can have profound effects on management decisions, as illustrated in an example for goldcrest in the Swedish breeding bird survey. In the example, if temporal random effects were not used, red listing would be highly influenced by the specific year in which it was done. When temporal random effects are used, red listing is stable over time. The methods are available in an R-package, `poptrend`.

Key-words: bird population, bird survey, generalized additive mixed model, generalized additive model, monitoring programme, population, population change, population management, population trend, trend

Introduction

Estimating trends of populations with large geographical distributions is a major but complex task in wildlife conservation (Balmford, Green & Jenkins 2003) and forms the basis for many broad-scale management decisions and classification schemes such as red listing according to the IUCN criteria (Rodrigues *et al.* 2006). The reliability of trend estimates and their associated uncertainties is therefore paramount (Akçakaya *et al.* 2000). Traditionally, trends have been modelled as linear changes in abundance, but it is now common to estimate nonlinear population trends (Link & Sauer 1998; Siriwardena *et al.* 1998; Fewster

*Correspondence author. E-mail: jonas.knape@slu.se

et al. 2000; Benton *et al.* 2002). Nonlinear trend estimates can reveal more information about population change than linear estimates. For instance, they may be more effective in picking up a recent decline in a population and be less sensitive to the time window over which the trend is computed. This is becoming ever more relevant as the time spans of monitoring programmes are increasing.

Because surveying populations across large spatial scales is costly, many large-scale monitoring programmes adopt fairly simple survey techniques using standardized counts of individuals repeated over time across a large number of spread out survey sites (Sutherland 2006). There is often large variation in counts among different localities, observers and other factors, as well as between visits (Gaston & McArdle 1994; Underwood 1994). Statistical methods need to be able to account for these sources of variation in order to reliably estimate uncertainty (Link & Sauer 1998). Standard methods for estimating trends often use site effects and continuous covariates in combination with a linear trend effect within a loglinear model framework to handle variation among survey localities and other factors (Gregory *et al.* 2005; Pannekoek & van Strien 2005; van Swaay *et al.* 2008).

Robust frameworks for estimating general nonlinear trends include generalized additive models (GAM; Wood 2006) and time-series models (Durbin & Koopman 2001; Fay & Punt 2013). GAMs can be seen as extensions of generalized linear models that enable the expected response to vary smoothly with continuous covariates. Fewster *et al.* (2000) developed a framework for the estimation of population trends using GAMs with a Poisson distribution for the response, site effects and a smooth component for time. The method attributes all temporal variation, short and long term, to the trend. With loglinear trend models, it is common to estimate both a linear long-term trend and short-term fluctuations in the form of annual indices (ter Braak *et al.* 1994). Such approaches have also been used in combination with smooth trends by first estimating annual indices and then smoothing the indices (Siriwardena *et al.* 1998). An alternative to this two-step approach is to use a model to decompose change into several components in order to more coherently handle uncertainty. Model-based approaches to decomposing nonlinear change include structural time-series models (Durbin & Koopman 2001) and wavelet analyses (Percival & Walden 2000), but usually are aimed at real valued data, as opposed to discrete count data. For count data, the range of methods available is more limited. GAMs, and their mixed model extensions (GAMM), provide tractable means to this end because the same machinery used to fit generalized linear mixed models can also be used to fit GAMMs (Wood 2006). Alternatively, GAMMs can be fit within a Bayesian framework. Other methods such as integrated nested Laplace approximations (INLA) applied to non-Gaussian structural time-series models could also be used (Rue, Martino & Chopin 2009).

Thomas (1996) identified four key components of count survey data in the context of estimating population change: prevailing trend, irregular perturbations, autocorrelation and sampling error. Here, I suggest using GAMMs to decompose overall population change into a long-term nonlinear smooth component representing the prevailing trend and a short-term random-effects component representing irregular perturbations. For computational reasons, I do not incorporate autocorrelation in the models but its effects on analyses are explored in simulations. The sampling error component is handled in a standard manner via a response distribution. The approach extends the methods developed by Fewster *et al.* (2000) to separate out irregular perturbations from the prevailing trend using the short-term random-effects component. GAMMs have been used previously to analyse population survey data (e.g. Mizuki *et al.* 2012; Ingersoll, Sewall & Amelon 2013; Melendez & Laiolo 2014), but I am not aware of any similar application of them to decompose population change into short- and long-term components. The models are applied to data on goldcrest *Regulus regulus* and greenfinch *Chloris chloris* from the Swedish Breeding Bird Survey and compared to the method of Fewster *et al.* (2000). Properties of the models are further investigated in simulations. I compare estimation with automatic selection of smoothing parameters, which is standard in the widely used `mgcv` package in R (Wood 2006), to estimation with fixed degrees of freedom (d.f.) as suggested by Fewster *et al.* (2000). An R-package `poptrend` based on `mgcv` is available.

Materials and methods

I will assume that the data consist of y_{it}, the number of individuals counted at site i at time t for $i = 1,\ldots,S$ and $t = 1,\ldots,T$. Hence, the data are an $S \times T$ matrix of counts, some of which may be missing. In addition, there may be covariates associated with time, sites and/or site by time. I will often refer to t as denoting year since this is the time scale used in many monitoring programmes, but in general, t could also represent some other time variable, for example month or week.

LINEAR TREND MODELS

Estimation of linear trends from count surveys is often done by loglinear regression (McCullagh & Nelder 1989). An often-used model from which trends may be derived is:

$$y_{it} \sim Po(\exp(a + bt + s_i)), \qquad \text{eqn 1}$$

where a is an intercept, b is the slope of the trend at the log-scale, and s_i are site effects to account for among-site variation (Pannekoek & van Strien 2005). The distribution of P-values using this model on simulated data with random among-year variation but no long-term change is heavily skewed towards the low end (Fig. 1). This implies that when among-year variation in abundance is fully random, in many cases there will be strong support for a linear trend over time in the data. This problem is primarily an issue resulting from incorrect estimates of uncertainty: the estimated slope of the trend will typically be small (Fig. 1).

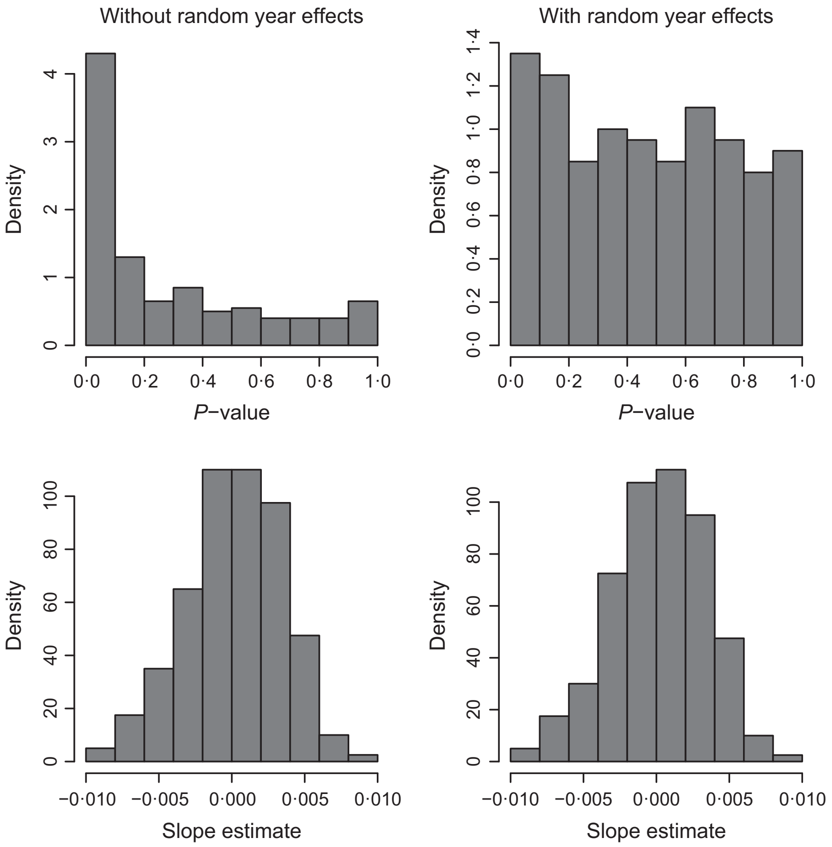

Fig. 1. Distribution of *P*-values (top panels) and estimates for the slope (bottom panels) of linear temporal trends from 200 simulated data sets using a Poisson distribution, each with 100 sites over 25 years and 15 sites with no systematic trend but random temporal variation without (left panels) and with (right panels) random year effects in the estimation model. The intercept was 0·8. Among-site and among-year variation were generated from normal distributions with standard deviation 0·3 and 0·2.

An alternative to model (1) is to estimate an index for each year using the model

$$y_{it} \sim Po(\exp(b_t + s_i)), \qquad \text{eqn 2}$$

where b_t is a fixed year effect. This type of model is incorporated in the software TRIM (Pannekoek & van Strien 2005) and frequently used in practice (e.g. Conrad *et al.* 2006).

Another alternative is to use a mixed model:

$$y_{it} \sim Po(\exp(a + bt + \varepsilon_t + s_i)), \quad \varepsilon_t \sim N(0, \sigma^2). \qquad \text{eqn 3}$$

The model uses year as both a continuous covariate and a random effect. Extended versions of this basic model have been applied to, for example, the North American Breeding Bird Survey data (Link & Sauer 2002; Sauer *et al.* 2008). In contrast to model (2), model (3) allows estimating a linear trend despite extra among-year variation, with the issue of inflated type I errors under random variation essentially removed (Fig. 1). In the next sections, I will extend trend estimation from generalized additive models (Fewster *et al.* 2000) to separate short-term fluctuations and long-term trends in the same way that the mixed model (3) extends the linear trend model (1).

NONLINEAR TREND MODELS

Early uses of smoothing methods for trend estimation applied smoothing to estimated population indices (Siriwardena *et al.* 1998), or separately to data from each surveyed site in the route regression approach used for the North American Breeding Bird Survey (James, McCulloch & Wiedenfeld 1996). Fewster *et al.* (2000) provided a framework for estimating nonlinear trends using GAM models, integrating the smoothing into a single model for the data. Their basic model is:

$$y_{it} \sim Po(\exp(a + S(t) + s_i)), \qquad \text{eqn 4}$$

where $S(t)$ is a smooth function of time representing (possibly) nonlinear changes in abundance. Unless one puts restrictions on the shape of the function $S(t)$, estimates from this model will be similar to estimates from model (2). The function $S(t)$ will essentially capture all temporal variability in the data. The d.f. therefore needs to be restricted. Fewster *et al.* (2000) recommended fixing the d.f. prior to analysis, using a third of the number of survey times, T, as a default value. Another method is to automatically select optimal d.f. via model selection or estimation (Wood 2011).

The GAM model (4) with restricted d.f. can be seen as an intermediate between the linear trend model (1) and the index model (2). Just as for the linear model (1), an assumption behind it is that the smooth function $S(t)$ is picking up all temporal variation except the variation provided by the Poisson. That is, errors from the smooth are assumed independent both within and among survey times. This could be viewed as a form of pseudoreplication (Hurlbert 1984) of the long-term trend and can give misleading estimates of uncertainty if there are additional short-term fluctuations in the data, as will be shown in a later section

A GAMM TREND MODEL

I propose separating smooth trend components from random among-year variation using the GAMM equivalent of the mixed effects model (3):

$$y_{it} \sim Po(\exp(a + S(t) + \varepsilon_t + s_i)),\\ \varepsilon_t \sim N(0, \sigma^2), \qquad \text{eqn 5}$$

where ε_t is a random effect for year. This model describes the temporal change as a combination of a smooth component, $S(t)$, and a random component, ε_t. To make the model identifiable, it is necessary to restrict the d.f. of the smooth component S, since an unpenalized smooth would be able to capture all among-year variation. With restrictions on the d.f., the S component of the model can be seen as a 'slowly' changing long-term trend component, while the random effect ε_t is capturing short-term, undirectional, fluctuations. If the sampling errors in a given year are not independent among sites, the short-term component may also consist of sampling variation. The exact meanings of 'long term' and 'short term' depend on the d.f. of the smooth and on the variance of the random year effect. The more restricted the smooth component is, the longer the time scale of the estimated trend and the larger the variance of the random effect. If there are no *a priori* reasons to select a specific degree of smoothing, the d.f. of the long-term trend as well as the variance of the short-term component can be determined automatically using either model selection approaches or by treating both components as random effects (Wood 2011).

For simplicity, I have presented all models with a Poisson response distribution, but since count data are often overdispersed with respect to the Poisson, other distributions, or quasi-distributions, can be used (ver Hoef & Boveng 2007).

OVERALL TRENDS

Model (5) describes the expected change in the population at each site *i*. To estimate overall change, total population size may be defined as the total number of individuals each year summed across all sites:

$$P(t) = \sum_i N_{it}$$

where N_{it} is the population size at site *i* in year *t*. From count data such as that discussed here, there is no direct access to the absolute population sizes N_{it}, but we can estimate them up to a proportionality factor c_{it} using the predicted counts. If observation errors are only due to missed detection of individuals (Kéry *et al.* 2009), the c_{it} may be interpreted as detection probabilities (see, e.g. Knape & Korner-Nievergelt 2015), but in general they give the proportions of the observed counts to the actual numbers of individuals. If we knew the value of $c_{it,}$ total population size in year *t* could be estimated as

$$P(t) = \frac{\sum_i N_{it}}{c_{it}} = \frac{\sum_i N_{it}^L \exp(\varepsilon_t)}{c_{it}},$$

where N_{it} is the predicted count given by $\exp(a + S(t) + \varepsilon_t + s_i)$ in eqn (5) using estimated values of parameters and random effects, and N_{it}^L is the predicted count when ignoring the temporal random effect (i.e. $N_{it}^L = \exp(a + S(t) + s_i)$. If we do not know c_{it}, we must assume that it has no trend over time or, if not, that we have covariates explaining its temporal variation. For now, we assume that $c_{it} = c_i$. We then define the long-term trend, standardized with respect to the first year as:

$$I_L(t) = P(t)/P(1)\exp(\varepsilon_1 - \varepsilon_t) = \sum_i (N_{it}^L/c_i)/\sum_i (N_{i1}^L/c_i)\\ = \exp(S(t) - S(1)), \qquad \text{eqn 6}$$

where $S(t)$ is the estimated smooth component representing the long-term trend. Here, the long-term trend is standardized with respect to the total predicted count at year 1 when random effects are ignored. Other standardizations are possible. A particularly useful alternative is to use the mean predicted count, that is $\sum_i (N_{i.}^L/c_i)$ where $N_{i.}^L$ is the average predicted count at site *i*, ignoring the temporal random effects. This has the benefit over the standardization in (6) that it is less affected by uncertainty in the estimate for the first year. Another alternative to reduce effects of uncertain data in the beginning of a survey is to use an interior year with more data as the reference (e.g. Fedy & Aldridge 2011).

The overall trend in eqn (6) is an estimate of the long-term component of population change. However, the short-term component represented by the random effect may also be of interest. To this end, we may define an index $I_S(t)$ incorporating the short-term fluctuations as:

$$I_S(t) = I_L(t)\exp(\varepsilon_t) = P(t)/P(1)\exp(\varepsilon_1)\\ = \sum_i (N_{it}/c_i)/\sum_i (N_{i1}^L/c_i) = \exp(S(t) - S(1) + \varepsilon_t), \qquad \text{eqn 7}$$

where ε_t is an estimate of the random effect in year *t*. The index is standardized with respect to the long-term component as in eqn (6) rather than with respect to the short-term index itself. Using the same standardization makes it possible to plot both the long-term trend $I_L(t)$ and the index $I_S(t)$ at the same scale.

In equations (6 and 7), the c_i cancel because of the assumption of temporal constancy, and because there are no terms varying with both time and space in the linear predictor of model (5), that is the trends are identical at all sites. If the proportionality constants vary systematically with time, the estimates for a specific year may over or under-represent the population size in the calculation of $I_L(t)$ and $I_S(t)$. Variation in the proportionality constant may arise, for example, from variation in sampling effort or detectability. For models where trends differ among sites (e.g. among habitats or regions), among-site variation in the proportionality constant can also be problematic because of the difficulty of appropriately weighting the changes between the different sites. In other cases, the sampling design may allow correction for the variation in the proportionality constant (Kéry *et al.* 2009), for example by including estimates of detection probabilities (Harrison *et al.* 2014) or the use of covariates believed to capture temporal or spatial variation in the proportionality constant. Studies on the UK Breeding Bird Survey have suggested that correcting for temporal variation in detection probabilities has little effect on trend estimates for most species (Newson *et al.* 2013).

Details on methods for computing confidence intervals for the trend estimates and changes in trends, treatment of covariates affecting sampling error or abundance, and how to check model fit are given in Appendix S1 (Supporting information).

SIMULATIONS

To compare the performance of the different methods for estimating nonlinear trends, I performed a set of simulations where a population is undergoing both long-term changes and short-term fluctuations. I considered three scenarios, each with 25 years of data, and 100 survey sites and Poisson-distributed sampling error: (i) no long-term trend, (ii) a long-term trend that is stable for the first 15 years, then declining to half the original level during approximately 5 years and then stable at this lower level for the remaining years, and (iii) a long-term trend that is first increasing and later declining. Log-scale site effects were simulated from a normal distribution with standard deviation 0·3 and log-scale random year effects from a normal distribution with standard deviation 0·1. For the third scenario, I ran an extra simulation where counts within the same sites were autocorrelated over time. This was done using a stationary AR(1) process on the log-scale with variance 0·2 and first-order coefficient 0·9. In all simulations, the average expected count in the beginning of the survey was 3.

For each of 500 simulated data sets for each scenario, I fitted smooth trend models with (5) and without (4) random year effects and with and without automatic selection of d.f., assuming a quasi-Poisson response (ver Hoef & Boveng 2007; the simulated data were not overdispersed). When fitting models with fixed d.f., I set the d.f. to one-third of the length of the time series. When fitting models with automatic selection of the d.f., I restricted the search so that the d.f. was forced to be less than a third of the length of the series. Models were fit using an R-package `poptrend` (url:github.com/jknape/poptrend), building on the package `mgcv` and its function gam, and using cubic regression splines for the long-term trend component (Wood 2006). For smoothing selection, generalized cross-validation was used. The way the function gam handles random effects means that variances of random effects were also estimated with generalized cross-validation instead of more standard maximum likelihood (ML) or restricted maximum likelihood estimation (REML; Wood 2011). Alternatively random effect variances and smoothing d.f. can be estimated with ML or REML by treating the smooth functions as random effects (Wood 2011). Uncertainty estimates were computed using parametric bootstrapping based on assuming asymptotic normality of the estimates (Harrison *et al.* 2014, Appendix S1). The code for the simulations is given in Appendix S2.

SWEDISH BREEDING BIRD SURVEY DATA

To test the GAMM trend model on real data, I estimated trends of two bird species using data from the Swedish Breeding Bird Survey. The survey is a monitoring scheme funded by the Swedish Environmental Protection Agency and run at Lund University to track the changes of common breeding birds in Sweden (Lindström *et al.* 2013; Green & Lindström 2015). The survey protocol used here consists of 716 quadrat routes laid out on a grid across Sweden. Each route is a 2 × 2 km square surveyed once a year during the breeding season by an observer counting all birds seen while walking along the 2 km sides of the square. In addition, point counts are performed along the routes, but we here only use the transect counts. I applied the GAMM trend estimation method to counts of goldcrest *Regulus regulus* and greenfinch *Chloris chloris* from the survey between 1998 and 2012. The goldcrest, a common breeding bird in Sweden, has been declining during the past decade and is included for the first time in the recently released version of the Swedish red list (Artdatabanken 2015). The greenfinch is another common breeding bird which has also been in sharp decline during the recent decade, attributed to trichomonosis (Lehikoinen *et al.* 2013), but is not included in the red list. While data are available for 1 km segments of the 2 × 2 km squares, species counts are generally low at each segment and the data were pooled from all segments on each route in each year. For some routes, some segments could never or not always be fully sampled, but the approximate proportion of segments covered is recorded. The proportion of the total 8-km distance covered was therefore used as an offset in the models (i.e. it was treated as a continuous covariate with coefficient fixed at 1). In some cases, this proportion was not recorded each year for a given route. Since the coverage proportions tend to be stable over time, I used the average recorded proportions in such years. I used latitude, age of the observer and day of the season as covariates, modelled as smooth terms (thin plate regression splines) with automatic selection of d.f. A binary covariate indicating whether the route was surveyed for the first time by an observer was also included (Kendall, Peterjohn & Sauer 1996). Site effects were treated as random to avoid estimating a separate site parameter for each route. Since counts may be overdispersed with respect to the Poisson distribution, a quasi-Poisson response model was used instead of the Poisson in eqns (2) and (3). Generalized cross-validation was used for smoothing selection and estimation of the variances of random effects (Wood 2006). R-code for the analysis can be found in Appendix S2.

Results

SIMULATIONS

On average across the simulations, all methods (with and without temporal random effects and automatic selection of d.f.) were able to reasonably recover the true deterministic trend in all three scenarios (scenario 3: Fig. 2, Appendices S5–S6; scenarios 1 and 2: Appendices S3 and S4). In scenarios 2 and 3, with changing populations, there was some tendency for the method with random year effects and automatic selection of d.f. to, on average, somewhat over-smooth the trend, producing a slight bias. If this is a concern, it could be remedied by inflating the estimated value of the d.f., but determining an appropriate amount of inflation for any specific scenario may require simulations. Mean-squared errors of the trend estimates with random year effects were comparable to, or slightly lower, than for the models without random year effects (Fig. 2, Appendices S3–S6). The main difference, however, was in the estimated uncertainty. For time periods when there was no or little true long-term change in the simulated populations, the models with random year effects indicated significant change (at the 10% level) at approximately the nominal rate. In contrast, for models fitted without random year effects, type I errors were inflated by a factor of almost 4 (Fig. 2, Appendices S3–S6). For instance, in scenario 3 around year 15, when the population is at its peak and there is no instantaneous change (no local decline or increase), models without

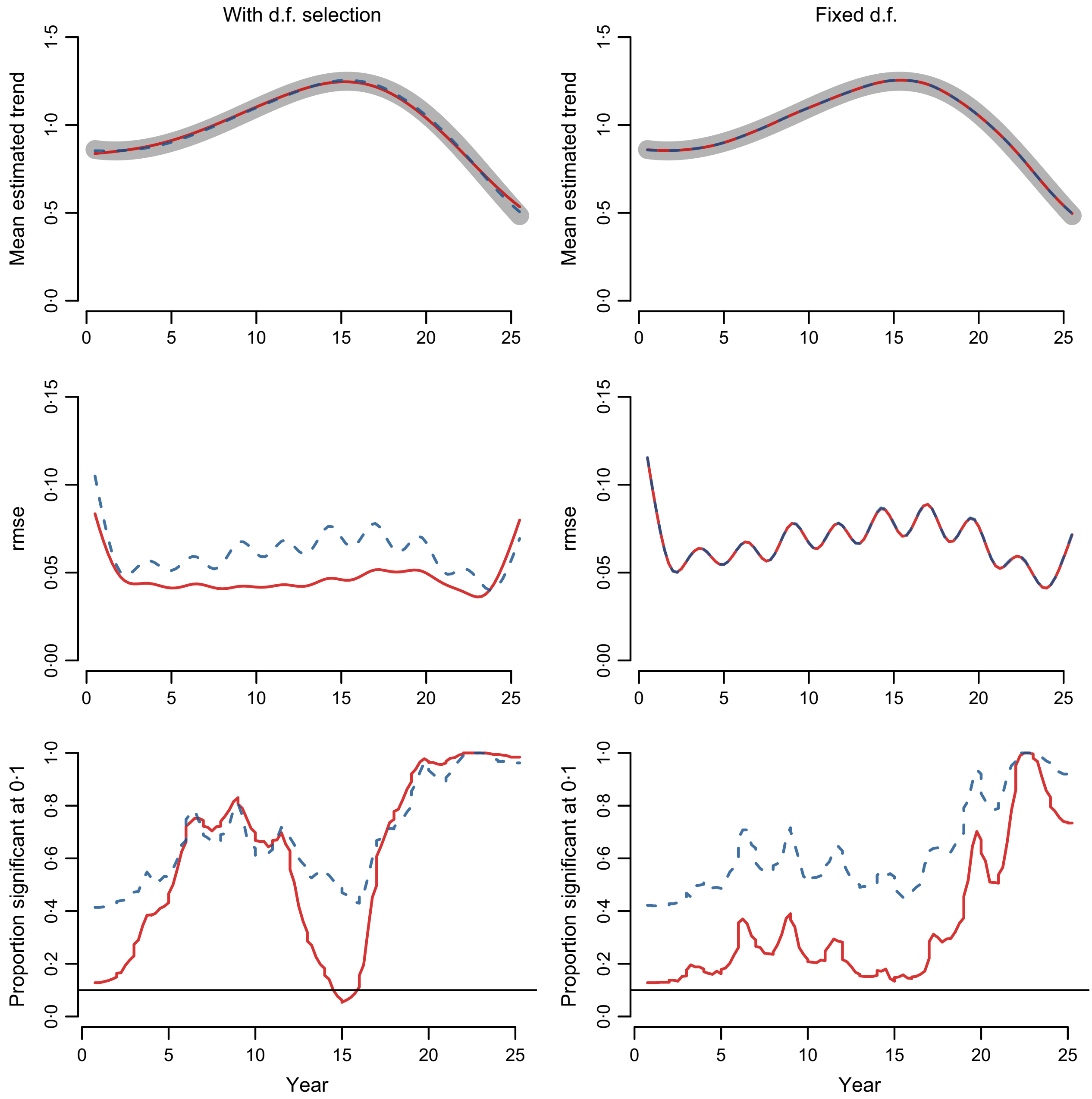

Fig. 2. Simulation results for a population with an initial increasing but then declining trend (scenario 3) with automatic selection of d.f. (left column) and with d.f. fixed to 9 (right column) for models with (solid lines) and without (dashed lines) random year effects. Top panels give the mean estimated trends (the long-term component only in the random year effects case) across 500 simulations; the thick grey shaded line is the true deterministic trend. Middle panels give root-mean-square errors (rmse) and lower panels the proportion of simulated trends that were significantly different from zero at a 10% alpha level according to quantiles of the bootstrap distribution. Similar plots for scenarios 1 and 2 can be found in Appendices S3 and S4.

random year effects indicated significant change in around 40% of the simulations while models with random year effects indicated significant change in the long-term component in ca. 10% of the simulations. This was the case both with and without automatic selection of d.f. During periods of true change, the proportion of simulations for which the change was significant was similar between models with and without random year effects when the d.f. were automatically selected (Fig. 2, Appendices S4–S6), indicating that the inclusion of random year effects did not have much effect on the power to detect long-term changes. Under fixed d.f., models with random year effects in some cases detected true changes less often than models without random year effects (Fig. 2). Reduced power to detect true long-term change when including temporal random effects is, however, to be expected also under automatic selection of d.f. in some cases.

One further observation from the simulations is that, for individual data sets, trends estimated from the models without random year effects are similar whether the d.f.

are fixed or automatically selected. For the models with random year effects, there is generally a difference (Appendices S3–S6). Because when fitting models with automatic selection of d.f. I restricted the d.f. to be less than the value used for the models fit with fixed d.f., the estimated d.f. tended to end up close to the maximum allowed value in the models without random year effects.

SWEDISH BREEDING BIRD SURVEY

Estimated trends for the goldcrest with automatic selection of d.f. with and without random year effects were qualitatively different (Fig. 3). The model with random year effects suggests a nearly linear decline in the long-term component and large short-term fluctuations. The model without random year effects in contrast suggests several smaller peaks and valleys, followed by a larger dip and then an almost full recovery of the population in the last few years. With fixed d.f., the estimated trends were more similar between the models with and without random year effects, but confidence intervals for the model with random year effects were much wider.

In the analysis of greenfinch data, all methods provided similar estimates of the trend, but with slightly wider confidence intervals for the models with random year effects (Fig. 4). For the greenfinch, these latter models suggest that the short-term fluctuations are relatively small.

For goldcrest, there were clear covariate effects of observer age and latitude, and for greenfinch, there was a clear effect of latitude and some effect of day of season, age and first survey (Appendix S7). Goodness-of-fit plots for these analyses display acceptable fits. Some deviation from normality can be seen in the estimated site effects, particularly for the greenfinch data (Appendix S7). This is unlikely to have a sizeable effect on the estimated trends (McCulloch, Searle & Neuhaus 2008). Estimated effects of covariates are given in Appendix S7. The trade-off between the d.f. of the long-term component and the variance of the random year effects is shown for the goldcrest data in fig. 5 in Appendix S7.

Discussion

Including temporal random effects in estimation of smooth trends makes it possible to separate long-term population change from short-term stochastic fluctuations. Simulations show that in the presence of random among-year variation in population size, including temporal random effects in GAM trend estimation greatly reduces spurious detection of long-term population change. Power to detect true changes in the long-term trend when including temporal random effects was similar to when not including them when using automatic selection of d.f. Some reduction in power is to be expected in general, for instance if the change in population size is small enough that the power of the model with random year effects is smaller than the baseline spurious detection rate of the GAM model without year effects (this rate

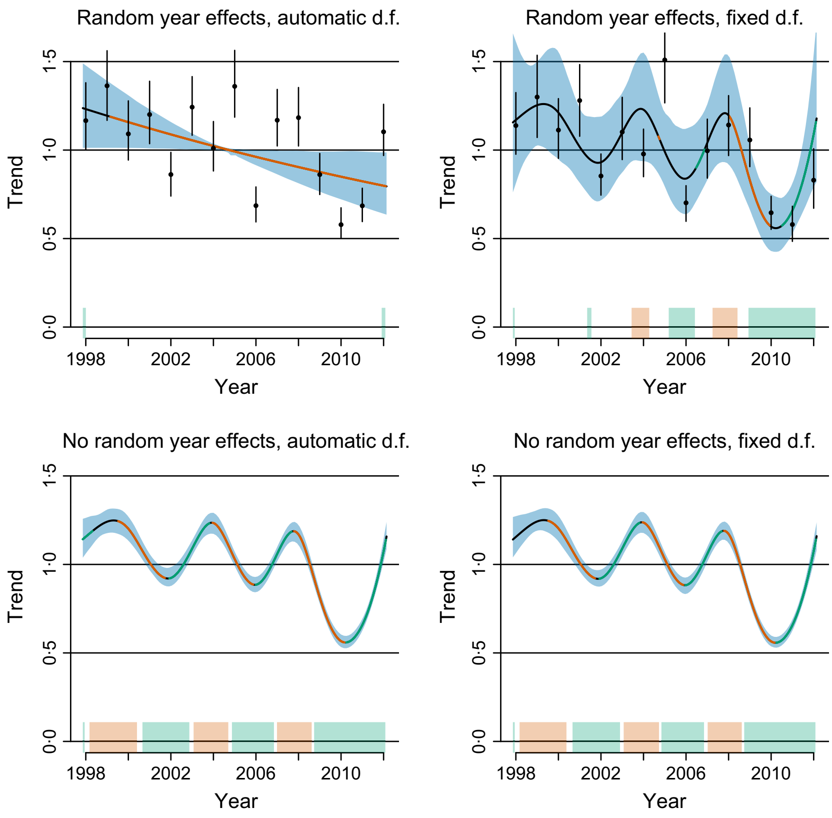

Fig. 3. Estimated trends for goldcrest data from the Swedish breeding bird survey. The top panels give estimates for models with random year effects, with automatic selection of d.f. (left) and d.f. fixed to 8 (right). The solid line in these panels is the estimated long-term component of the trend, while the points are estimates of the trend with estimates of the random year effects superimposed. The lower panels give estimates of the trend for models without random year effects. The estimates are standardized with respect to the mean of the long-term component of the trends. Confidence intervals (shaded area and vertical lines) are computed from the 2·5% and 97·5% quantiles of the bootstrap distributions. For periods where there is a significant increase or decrease in the trend at the 5% level, the trend line is coloured, respectively, in orange and green. Periods where the curvature is significantly positive or negative are marked by green and orange rectangles at the bottom of the panels (Appendix S1).

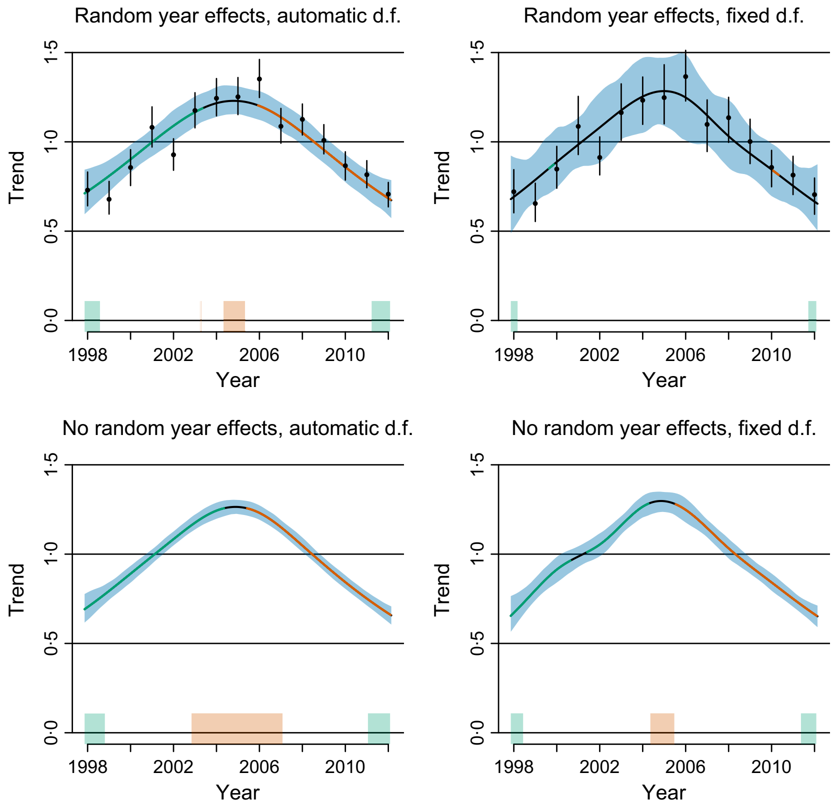

Fig. 4. Estimated trends for greenfinch data from the Swedish breeding bird survey. For explanation of details, see caption for Fig. 3.

was around 40% in the simulations). The GAMM trend estimation had similar or better properties in terms of mean square error, particularly under automatic selection of d.f. Furthermore, the GAMM framework allows flexible modelling of effects of covariates, including smooth effects of continuous covariates, or random effects for categorical covariates (Ingersoll, Sewall & Amelon 2013).

The exact interpretations of the short-term and long-term components depend on the d.f. of the long-term component and on how the temporal random effect is modelled. Here, I have used independent random effects for each year, but in principle an autocorrelated process, such as an autoregressive random effect, could be used instead. For survey data sets with short time series, such as the Swedish breeding bird survey, there may not be enough temporal replication to estimate an autoregressive parameter, but it might be an alternative for surveys with longer time series.

MANAGEMENT IMPLICATIONS

Although it is notoriously difficult to estimate variances in population size (Gaston & McArdle 1994), it is known that population abundance can fluctuate considerably among years (Underwood 1994). Fluctuations may be due to environmental stochasticity, population regulation, migration or other factors (Gaston & McArdle 1994). Another possible cause of erratic fluctuations is that measurement/sampling errors may not be independent among sampled sites in a given year which may cause apparent but non-demographic among-year variation in population size (Knape, Besbeas & de Valpine 2013). Data on goldcrest from the Swedish Breeding Bird Survey analysed here suggest strong, apparently random, among-year variation. The separation of fluctuations into a long-term trend component and short-term fluctuations can have consequences for the interpretation of trends. In a conservation setting, for instance, do we need to react to a population declining by 30% from one year to the next, if the change is in line with the normal among-year variation in the population? The goldcrest analysis provides an example of the different conclusions that could be reached by the two respective models. Swedish bird species are red listed partly based on the change in the population during the last 10 years, as estimated from the survey (using log-linear models). Suppose that the time period on which red listing of the goldcrest was based on was 2002–2012, and that smooth trend models without random year effects were used for red list classification. The estimates without random year effects would suggest an increase in the population of 16% with a 95% confidence interval of (7%, 27%). Had the time period for red listing instead been 2001–2011, the trend would suggest a strong population decline of −35% (−40%, −30%). The estimated change in the population therefore goes from a moderate increase to a strong decline by shifting the time period by 1 year, possibly leading to the goldcrest being included in a different red list category. For the proposed method with random year effects, the decline in the long-term component between 2002 and 2012 is −27% (−45%, −2%), and for 2001–2011, the decline is almost identical, −27% (−45%, −2%), due to the approximate linearity of the estimated trend. The difference between the methods is caused by the former model picking up higher counts in 2011 and 2012 as evidence of an increasing trend (Fig. 3). Effectively, it is picking up among-year variation in the trend that is considered short-term fluctuations by the random-effects model. In contrast, for the greenfinch data, most of the among-year variation was composed of long-term changes and models with and without temporal random effects gave similar results.

GUIDANCE FOR APPLIED USES

If the goal of analysis is an estimate of population trend, and estimates of statistical uncertainty are not a priority, then random year effects may not necessarily be needed in GAM analyses; both models provided effectively unbiased estimates in the simulations. If automatic selection of d.f. is used, trends estimated without random year effects would, however, typically be more variable (less smooth) than the corresponding long-term estimate with random year effects. Separating out long-term trend components from short-term fluctuations can reduce noise in management decisions, as demonstrated in the analysis of the goldcrest population in Sweden above.

In management settings, accurate estimation of uncertainty is vital for reliably weighting evidence. The main benefit of including random year effects is more accurate estimates of uncertainty of the long-term trend. This is shown in the simulations where the false error rate is around 40% when there is no long-term change but stochastic fluctuations in the population. Hence, if long-term change (prevailing trend) and associated uncertainty is the focus of analysis, I recommend the inclusion of random year effects. This will often result in more uncertain estimates of the trend, but the uncertainty estimates are more realistic concerning the long-term trend. Sample size requirements may therefore be higher when random year effects are used. In particular, longer time series and more accurate annual indices may be necessary. It is difficult to give general guidelines to how large samples will be needed because many factors contribute in different ways to the precision of the estimates. To check the performance of the proposed method in specific settings not covered by the simulations here, or to assess power, simulations using conditions pertaining to the particular survey are advised. The R-script provided in Appendix S2 could serve as a starting point for such analyses.

An important choice faced by the analyst is whether or not to use automatic selection of d.f. If random year effects are not included, simulations here suggest there will often be little practical difference. On the other hand, when random year effects are included the difference can be large. Sometimes, report requirements or considerations about the temporal scale of the inference of interest will provide guidance to fixing the d.f. to some pre-determined value in order to achieve trends with the desired amount of smoothing (Fewster *et al.* 2000). In the absence of *a priori* requirements on the d.f., a statistically optimal trade-off between the d.f. of the smooth and the variance of the random effects can be found. This may be done using a model selection procedure such as cross-validation, or, alternatively, by considering the smooth components as random effects and estimate associated variance parameters via standard procedures (Wood 2011). Simulations in this paper suggest that automatic selection of d.f. performs well in recovering long-term trends.

Modern software, such as the R-package `mgcv`, allows these methods to be relatively easily applied, even for larger surveys. To further aid researchers and practitioners in adopting GAMM methods for trend estimation, an R-package, `poptrend`, is available (url:github.com/jknape/poptrend).

Acknowledgements

I am grateful to Debora Arlt, Alison Johnston, Åke Lindström and Marc Kéry whose comments greatly improved the manuscript and to Michal Zmihorski for input on the R-code. I thank everyone collecting data for the Swedish Breeding Bird survey, which is supported by the Swedish Environmental Protection Agency and all 21 regional county boards of Sweden. This work was funded by grant 621-2012-4076 from the Swedish Research Council VR.

Data accessibility

Survey data for goldcrest and greenfinch: uploaded as online supporting information.

R scripts: uploaded as online supporting information.

References

Akçakaya, H.R., Ferson, S., Burgman, M.A., Keith, D.A., Mace, G.M. & Todd, C.R. (2000) Making consistent IUCN classifications under uncertainty. *Conservation Biology*, **14**, 1001–1013.

Artdatabanken (2015) *Rödlistade Arter i Sverige 2015*. Artdatabanken SLU, Uppsala.

Balmford, A., Green, R.E. & Jenkins, M. (2003) Measuring the changing state of nature. *Trends in Ecology & Evolution*, **18**, 326–330.

Benton, T.G., Bryant, D.M., Cole, L. & Crick, H.Q.P. (2002) Linking agricultural practice to insect and bird populations: a historical study over three decades. *Journal of Applied Ecology*, **39**, 673–687.

ter Braak, C.J.F., van Strien, A.J., Meijer, R. & Verstrael, T.J. (1994) Analysis of monitoring data with many missing values: which method. *Bird*, **1992**, 663–673.

Conrad, K.F., Warren, M.S., Fox, R., Parsons, M.S. & Woiwod, I.P. (2006) Rapid declines of common, widespread British moths provide evidence of an insect biodiversity crisis. *Biological Conservation*, **132**, 279–291.

Durbin, J. & Koopman, S.J. (2001) *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford.

Fay, G. & Punt, A.E. (2013) Methods for estimating spatial trends in Steller sea lion pup production using the Kalman filter. *Ecological Applications*, **23**, 1455–1474.

Fedy, B.C. & Aldridge, C.L. (2011) The importance of within-year repeated counts and the influence of scale on long-term monitoring of sage-grouse. *The Journal of Wildlife Management*, **75**, 1022–1033.

Fewster, R.M., Buckland, S.T., Siriwardena, G.M., Baillie, S.R. & Wilson, J.D. (2000) Analysis of population trends for farmland birds using generalized additive models. *Ecology*, **81**, 1970–1984.

Gaston, K.J. & McArdle, B.H. (1994) The temporal variability of animal abundances: measures, methods and patterns. *Philosophical Transactions: Biological Sciences*, **345**, 335–358.
Green, M. & Lindström, Å. (2015) Övervakning av fåglarnas populationsutveckling. Årsrapport för 2014. – Rapport, Biologiska institutionen, Lunds Universitet. 86 pp.
Gregory, R.D., van Strien, A., Vorisek, P., Gmelig Meyling, A.W., Noble, D.G., Foppen, R.P.B. & Gibbons, D.W. (2005) Developing indicators for European birds. *Philosophical Transactions: Biological Sciences*, **360**, 269–288.
Harrison, P.J., Buckland, S.T., Yuan, Y., Elston, D.A., Brewer, M.J., Johnston, A. & 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**, 1650–1660.
ver Hoef, J.M. & Boveng, P.L. (2007) Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology*, **88**, 2766–2772.
Hurlbert, S.H. (1984) Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, **54**, 187–211.
Ingersoll, T.E., Sewall, B.J. & Amelon, S.K. (2013) Improved analysis of long-term monitoring data demonstrates marked regional declines of bat populations in the eastern United States. *PLoS ONE*, **8**, e65907.
James, F.C., McCulloch, C.E. & Wiedenfeld, D.A. (1996) New approaches to the analysis of population trends in land birds. *Ecology*, **77**, 13–27.
Kendall, W.L., Peterjohn, B.G. & Sauer, J.R. (1996) First-time observer effects in the North American Breeding Bird Survey. *The Auk*, **113**, 823–829.
Kéry, M., Dorazio, R.M., Soldaat, L., van Strien, A., Zuiderwijk, A. & Royle, J.A. (2009) Trend estimation in populations with imperfect detection. *Journal of Applied Ecology*, **46**, 1163–1172.
Knape, J., Besbeas, P. & de Valpine, P. (2013) Using uncertainty estimates in analyses of population time series. *Ecology*, **94**, 2097–2107.
Knape, J. & Korner-Nievergelt, F. (2015) Estimates from non-replicated population surveys rely on critical assumptions. *Methods in Ecology and Evolution*, **6**, 298–306.
Lehikoinen, A., Lehikoinen, E., Valkama, J., Väisänen, R.A. & Isomursu, M. (2013) Impacts of trichomonosis epidemics on Greenfinch *Chloris chloris* and Chaffinch *Fringilla coelebs* populations in Finland. *Ibis*, **155**, 357–366.
Lindström, Å., Green, M., Paulson, G., Smith, H.G. & Devictor, H. (2013) Rapid changes in bird community composition at multiple spatial scales in response to recent climate change. *Ecography*, **36**, 313–322.
Link, W.A. & Sauer, J.R. (1998) Estimating population change from count data: application to the North American breeding bird survey. *Ecological Applications*, **8**, 258–268.
Link, W.A. & Sauer, J.R. (2002) A hierarchical analysis of population change with application to Cerulean warblers. *Ecology*, **83**, 2832–2840.
McCullagh, P. & Nelder, J.A. (1989) *Generalized Linear Models*. CRC Press, Boca Raton, FL, USA.
McCulloch, C.E., Searle, S.R. & Neuhaus, J.M. (2008) *Generalized Linear and Mixed Models*, 2nd edn. Wiley, Hoboken, New Jersey, USA.
Melendez, L. & Laiolo, P. (2014) The role of climate in constraining the elevational range of the Water Pipit *Anthus spinoletta* in an alpine environment. *Ibis*, **156**, 276–287.
Mizuki, I., Sakaguchi, S., Fukushima, K., Sakai, M., Takayanagi, A., Fujiki, D. & Yamasaki, M. (2012) Among-year variation in deer population density index estimated from road count surveys. *Journal of Forest Research*, **18**, 491–497.
Newson, S.E., Massimino, D., Johnston, A., Baillie, S.R. & Pearce-Higgins, J.W. (2013) Should we account for detectability in population trends? *Bird Study*, **60**, 384–390.
Pannekoek, J. & van Strien, A. (2005) *TRIM 3 Manual (Trends & Indices for Monitoring Data)*. Statistics Netherlands, Voorburg, the Netherlands.
Percival, D.B. & Walden, A.T. (2000) *Wavelet Methods for Time Series Analysis*. Cambridge University Press, New York, NY, USA.
Rodrigues, A.S.L., Pilgrim, J.D., Lamoreux, J.F., Hoffmann, M. & Brooks, T.M. (2006) The value of the IUCN Red List for conservation. *Trends in Ecology & Evolution*, **21**, 71–76.
Rue, H., Martino, S. & Chopin, N. (2009) Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society B*, **71**, 319–392.
Sauer, J.R., Link, W.A., Kendall, W.L., Kelley, J.R. & Niven, D.K. (2008) A hierarchical model for estimating change in American woodcock populations. *Journal of Wildlife Management*, **72**, 204–214.
Siriwardena, G.M., Baillie, S.R., Buckland, S.T., Fewster, R.M., Marchant, J.H. & Wilson, J.D. (1998) Trends in the abundance of farmland birds: a quantitative comparison of smoothed common birds census indices. *Journal of Applied Ecology*, **35**, 24–43.
Sutherland, W.J., ed. (2006) *Ecological Census Techniques: A Handbook*, 2nd edn. Cambridge University Press, Cambridge, UK.
van Swaay, C.A.M., Nowicki, P., Settele, J. & van Strien, A.J. (2008) Butterfly monitoring in Europe: methods, applications and perspectives. *Biodiversity and Conservation*, **17**, 3455–3469.
Thomas, L. (1996) Monitoring long-term population change: why are there so many analysis methods? *Ecology*, **77**, 49–58.
Underwood, A.J. (1994) On beyond BACI: sampling designs that might reliably detect environmental disturbances. *Ecological Applications*, **4**, 4–15.
Wood, S. (2006) *Generalized Additive Models: An Introduction With R*. CRC Press, Boca Raton, FL, USA.
Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**, 3–36.

Received 11 November 2015; accepted 8 June 2016
Handling Editor: Gavin Siriwardena

Supporting Information

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

Appendix S1. Confidence intervals, covariates and model checking.

Appendix S2. R-code and data.

Appendix S3. Simulation study scenario 1.

Appendix S4. Simulation study scenario 2.

Appendix S5. Simulation study scenario 3.

Appendix S6. Simulation study scenario 3 with autocorrelation.

Appendix S7. Additional figures for the SBBS analysis.