

LETTER

Estimating the annual number of breeding attempts from breeding dates using mixture models

Thomas Cornulier,^{1,2*} David A. Elston,³ Peter Arcese,⁴ Tim G. Benton,⁵ David J.T. Douglas,⁶ Xavier Lambin,¹ Jane Reid,¹ Robert A. Robinson⁷ and William J. Sutherland⁸

Abstract

Well-established statistical methods exist to estimate variation in a number of key demographic rates from field data, including life-history transition probabilities and reproductive success per attempt. However, our understanding of the processes underlying population change remains incomplete without knowing the number of reproductive attempts individuals make annually; this is a key demographic rate for which we have no satisfactory method of estimating. Using census data to estimate this parameter from requires disaggregating the overlying temporal distributions of first and subsequent breeding attempts. We describe a Bayesian mixture method to estimate the annual number of reproductive attempts from field data to provide a new tool for demographic inference. We validate our method using comprehensive data on individually-marked song sparrows *Melospiza melodia*, and then apply it to more typical nest record data collected over 45 years on yellowhammers *Emberiza citrinella*. We illustrate the utility of our method by testing, and rejecting, the hypothesis that declines in UK yellowhammer populations have occurred concurrently with declines in annual breeding frequency.

Keywords

Bayesian inference, conservation ecology, demographic parameters, life-history variation, population dynamics, reproductive success.

Ecology Letters (2009) 12: 1–10

INTRODUCTION

A key aim in ecology is to predict reliably the spatial and temporal dynamics of natural populations. Such understanding requires detailed knowledge of patterns of variation in several key demographic rates, especially survival, dispersal and the frequency and success of reproductive attempts by individual animals (Pease & Grzybowski 1995; Lande *et al.* 2003). However, it is difficult, in practice, to collect sufficient demographic data to allow detailed

estimation of the demographic rates. Indeed, a major, recent advance in population ecology has been the development and application of rigorous statistical methods to estimate demographic rates from incomplete or uncertain empirical monitoring data already available for many species worldwide. For example, sophisticated capture–mark–recapture models are now used to estimate apparent survival, migration and life-history transition probabilities from individual encounter histories, for species where only a fraction of all individuals can be observed in any given

¹UKPopNet, Institute of Biological and Environmental Sciences, Zoology Building, Tillydrone Avenue, University of Aberdeen, Aberdeen AB24 2TZ, UK

²Centre d'Etudes Biologiques de Chizé, CNRS, Villiers en Bois, 79360 Beauvoir sur Niort, France

³Biomathematics & Statistics Scotland, Craigiebuckler, Aberdeen AB15 8QH, UK

⁴Centre for Applied Conservation Research, Forest Sciences, 3rd Floor, 2424 Main Mall, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

⁵Institute of Integrative & Comparative Biology, Miall Building, University of Leeds, Leeds LS2 9JT, UK

⁶RSPB Scotland, Dunedin House, 25 Ravelston Terrace, Edinburgh EH4 3TP, UK

⁷British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU, UK

⁸Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK

*Correspondence: E-mail: cornulier@cebc.cnrs.fr

census (Lebreton *et al.* 1992; Nichols *et al.* 2005). Similar probabilistic frameworks have been used to estimate the success of individual reproductive attempts (Mayfield 1961; Johnson 1979; Aebischer 1999). However, there is, as yet, no satisfactory method to determine the number of reproductive attempts that individuals make annually, except where these can be directly counted by the continuous observation of marked individuals (Thompson *et al.* 2001). Because variation in breeding frequency can greatly influence an individual's total seasonal reproductive output and population growth rate (Pease & Grzybowski 1995; Wilson & Arcese 2003), the lack of an appropriate method to estimate breeding frequency currently precludes the comprehensive analysis of demographic change in many free-living species (Siriwardena *et al.* 2000).

Direct measurement of the number of reproductive attempts that individuals make within a season requires continuous monitoring of a sample of marked animals in nature – a goal that is, at best, time-consuming but often impossible. Accurate data on individual breeding frequency are therefore available from a relatively few intensively studied species within restricted study areas (e.g. Thomson & Cotton 1999; Hansson *et al.* 2000; Browne & Aebischer 2004; Nagy & Holmes 2005; Weggler 2006). In contrast, incomplete data on breeding frequency are often collected opportunistically through census-based studies. Such monitoring programmes may collect data in large quantities. For example, in the UK, each year amateur ornithologists locate about 30 000 nests of a range of bird species, which is used to report estimated annual laying dates (Crick *et al.* 2003). Methods able to exploit such opportunistic data sets to estimate the mean number of breeding attempts per individual per season have the potential to improve substantially the reliability of analyses aimed at identifying key demographic factors affecting population growth rate. Here, we propose, and then test, a statistical method designed to estimate the mean number of breeding attempts made by individuals each year from routinely collected data on the population-wide, temporal distribution of reproductive attempts.

Robust estimates of the annual number of reproductive attempts per individual should greatly improve our ability to predict population trends, as we will then have access to methods that predict survival, migration, reproductive output per attempt and the number of attempts. This inability to estimate all parameters has handicapped us to date. For example, despite considerable research, the underlying cause of recent declines in many western European populations of farmland birds remains uncertain (Donald *et al.* 2001; Newton 2004). One key hypothesis is that these population declines are due to a reduction in the mean number of reproductive attempts made by individuals each season (Siriwardena *et al.* 2000), perhaps due to earlier harvesting of crops (Donald *et al.* 2002). This hypothesis

remains largely untested, however (but see Browne & Aebischer 2004), due to a lack of comprehensive data on temporal and spatial variation in individual breeding frequency for a representative sample of species.

Here, we present a new Bayesian mixture method to estimate breeding frequency, defined as the average number of breeding attempts made per individual per season, from data describing the temporal distribution of breeding attempts observed at the population level. We assess the performance of the method using data on individual breeding frequency from an intensive 30-year study of song sparrows *Melospiza melodia*, where the phenology and breeding history of over 300 individually marked females has been accurately documented (Smith 2006). We then apply our method to an opportunistically collected data set describing the temporal distribution of breeding attempts in yellowhammers *Emberiza citrinella*, a European farmland bird that has experienced substantial population decline. By illustrating the estimation of temporal variation in breeding frequency for a species where comprehensive data from marked individuals are unavailable, we demonstrate how this is a valuable new tool for demographic inference. We demonstrate its utility by testing the hypothesis that declines of farmland birds in the UK are a consequence of reductions in the mean annual number of reproductive attempts by individual breeding birds.

MATERIAL AND METHODS

Principle of the method

Our aim is to estimate the mean number of breeding attempts made per individual per year, hereafter B , from the temporal distribution of breeding dates. In the absence of a direct measure of mean breeding frequency of N breeding pairs, this information can be estimated from the proportion, p_1 , of observed breeding attempts that are an individual pair's first breeding attempt of the season. This is because, by definition, the total number of breeding attempts (N_t) is the sum of the number of first breeding attempts ($N_1 = N$, because by definition each breeding pair must have a first breeding attempt) plus the number of subsequent breeding attempts (N_{2+}). For example, if $N = 20$ breeding pairs attempted $N_t = 60$ broods in one year, then $N_1 = 20$ were first attempts and $N_{2+} = 40$ were subsequent attempts. The average number of breeding attempts per pair is $(60/20) = 3$, which can in general be derived from the formula $B = (N_t/N) = (N_t/N_1) = 1/p_1$.

The estimation of p_1 from available data is therefore the statistical challenge that needs to be solved. To do so, we used a 'mixture model' (Congdon 2001), a statistical method that allows the properties of mixtures of overlapping distributions to be described quantitatively.

Data and model

We now consider the task of estimating p_1 from the temporal distribution of observed breeding dates for a population of interest. We initially made three assumptions regarding the population-wide distribution of breeding dates within any one year.

Assumption 1

The distribution of first breeding dates is sufficiently distinct from the distribution of breeding dates of subsequent broods to enable a two-component mixture model to be satisfactorily fit to the distribution of all observed breeding dates. This enables the proportion of observed breeding attempts that comprise first (p_1) and subsequent ($1 - p_1$) attempts to be estimated, as required to estimate B . The two components must reflect a population-wide pattern, not different breeding cohorts.

Assumption 2

The distribution of breeding dates of the first attempts can reasonably be approximated by a skew-normal distribution with location, scale and shape parameters to be estimated from data. The flexibility of this distribution, compared to the standard normal distribution, should accommodate common situations, for example, where most pairs breed almost as soon as climatic constraints permit combined with an asymmetrically heavy right tail of relatively late breeders (Wegglar 2006). Other standard distributions, such as the gamma, would equally fit right skewness with three parameters (to allow for the arbitrary choice of zero on the calendar day scale) but lack flexibility to potential left-skewed data.

Assumption 3

The breeding dates of all subsequent attempts follow a standard normal distribution with mean and variance parameters to be estimated from the data. The use of the standard normal distribution can be justified here from the multiplicity of influences on subsequent breeding dates, including failures at varying times of first broods and multiple repeat broods.

Writing the probability density function of the skewed normal distribution with location μ_1 , scale σ_1 and shape parameter α as $\text{SN}(X, \mu_1, \sigma_1, \alpha)$, and writing the probability density function of the standard normal distribution with mean μ_2 and standard deviation σ_2 as $N(X, \mu_2, \sigma_2)$, then the probability density function of a mixture containing a proportion p_1 from the first distribution is given by

$$f(X) = p_1 \text{SN}(X, \mu_1, \sigma_1, \alpha) + (1 - p_1) N(X, \mu_2, \sigma_2) \quad (1)$$

Because we wish to ensure $\mu_2 > \mu_1$, we later find it convenient to write eqn 1 using $\mu_2 = \mu_1 + \delta$, with $\delta > 0$.

Model estimation

The above model can be relatively easily fitted to available data on population-wide distributions of observed breeding dates within a Bayesian framework using Markov chain Monte Carlo (MCMC) methods. In this study we are concerned with historical data coming from long-term nest record schemes that present highly variable amounts of information across years. The Bayesian hierarchical approach enables the efficient use of sparse data by sharing information from different sources (Congdon 2001). Sharing information can be achieved in two ways. First, by incorporating random effects that smooth model parameters over a number of groups, which could be species, years, populations or classes of individuals within a population for instance. Including year as a random effect enabled us to obtain year-specific estimates of p_1 from all years contained within multi-year data sets. Such use of random effects modelling is particularly useful in a mixture distribution context where data quantity or quality varies among years, as for some years, data may be sparse or the combined distribution may show little separation into the components. In either case, it would not be possible to fit the model to such years individually. Second, the Bayesian method allows formal inclusion of prior information, for example parameter estimates from previous studies on other populations of the same or similar species.

When fitting eqn 1 to multi-year data sets of observed breeding dates, four of the six parameters were allowed to vary between years. We found that the skewness parameter could not be sufficiently informed from the data in many years due to small sample sizes and so was considered to be fixed. The year-specific model can thus be written as:

$$f(X) = p_{1i} \text{SN}(X, \mu_{1i}, \sigma_{1i}, \alpha) + (1 - p_{1i}) N(X, \mu_{1i} + \delta, \sigma_{2i}) \quad (2)$$

Parameters allowed to take year-specific values are denoted by the subscript i . The year-specific mixing proportions p_{1i} were assumed to follow a Normal distribution on the logit scale, thus: $\text{logit}(p_{1i}) \sim N(m_p, s_p)$. The differences $\mu_{2i} - \mu_{1i} = \delta$ were assumed to be constant across years.

Given our Bayesian framework, a final step in model building requires declaring priors for the underlying fixed parameters and the parameters describing distributions of random effects. These were chosen to be uninformative bounded within the limits of realistic parameter values.

In order to estimate the parameters of the breeding date distributions and the proportion p_1 of first and subsequent broods, we fitted our model by MCMC simulation using OpenBUGS 2.2.0 (BUGS code given in Appendix S1 in Supporting Information) called from R's BRugs library <http://mathstat.helsinki.fi/openbugs/> (Thomas *et al.* 2006;

R Development Core Team 2007). We used 500 000 updates of the model, after discarding a burn-in set of 5000. To reduce autocorrelation in the MCMC chain, one in 100 updates was kept to form the posterior sample.

Application when all laying dates are observed

To validate our mixture model method for estimating the average number of breeding attempts made per individual per year from distributions of observed laying dates, we used long-term life-history data from a song sparrow population resident on Mandarte Island, British Columbia, Canada. This small (4–74 breeding pairs) population has been subject to continuous study since 1975 (Smith 2006), when all population members were initially colour-ringed and, thereafter, the number and date of all reproductive attempts made annually by each breeding female were comprehensively documented, as described in Smith (2006). Briefly, all occupied territories were visited annually each 3–5 days from March to August, and all nests were located and monitored to fledging or failure. Laying dates were determined directly for nests located before clutch completion or back-calculated precisely by observing hatching dates, chick size and development for nests found subsequently. For our analyses here, we used the number of breeding attempts per female, because pair-based estimates underestimate the true number of breeding attempts per season when females mate with more than one territorial male in a season. Excluding 5.7% of nests for which laying dates were not determined exactly left us with 2532 nests (from 11 to 179 per year). First and last breeding, breeding season length and the number of attempts made by each individual varied markedly among years, reflecting variation in climate and population density (Wilson & Arcese 2003; Smith 2006). We fitted the mixture model to the observed distribution of laying dates across all years of the study after centring by year, estimated p_1 and hence B , and then compared estimated values B_{est} to the true observed values B_{obs} .

To validate our assumptions in the case of the yellowhammer, we used data collected during May to August 2005–2006 as part of a short-term, intensive study at five sites in Aberdeenshire, Scotland (Douglas *et al.* 2009). All nests were located by mapping territorial males and observing breeding behaviour. Nests containing eggs were visited about every 3 days to determine hatch date. Individual pairs were identified by colour rings, plumage characteristics or by their spatial isolation from other pairs. First egg date was calculated by extrapolating back using the age of the brood, by observing the hatching date or by comparison to broods of known age, assuming that one egg was laid each day, that incubation commenced upon completion of the clutch and lasted 13 days. These data allowed us to estimate directly the number of breeding

attempts made per individual per year (B_{obs}) and to examine the degree to which overall distributions of laying dates fulfilled assumptions 1–3 of our mixture model. Sample sizes were however too small for us to attempt fitting the mixture model to these data.

Effect of sample size on estimates of breeding frequency

We assessed the properties of the mixture model estimate of p_1 in relation to sample size using simulated data sets with total sample size N varying from 400 to as few as 25 nest records, in order to exclude other sources of variation, such as the violation of model assumptions. We drew laying dates from two distributions with fixed parameters: (1) a skew-normal distribution simulating first reproductive attempts with location = -35, scale = 20 and skewness = 2; (2) a normal distribution simulating all subsequent reproductive attempts with location = 35 and scale = 20. The target proportion of first reproductive attempts was set to $p_1 = 0.6$, corresponding to an average of 1.67 attempts per breeding pair. Each data item was drawn from distribution (1) with probability p_1 and from (2) otherwise (Fig. S2.1 in Appendix S2). For each sample size N , we simulated 100 data sets and estimated the parameters and mixing proportion of the underlying distributions, using the same model as for the song sparrow and yellowhammer (BTO, UK-wide) data sets, albeit without random effects associated with year.

Application when only a sample of laying dates is observed

After validating our method and assumptions across data sets for which B was observed directly, we then applied it to estimate B using a large, multi-year data set of observed laying dates for yellowhammers where B was not observed directly. These data comprised *ad hoc* observations of first egg dates for yellowhammers obtained from the British Trust for Ornithology's Nest Record Scheme, covering 1960–2004. As nests are typically revisited at intervals of several days, first egg dates can be estimated within a range of calculated minimum and maximum possible dates (for a comprehensive account of the Nest Record Scheme, see Crick *et al.* 2003). For simplicity, we used the mid-point of this interval as the estimated laying date of each nest. The analysis was based on 6286 nests (from 36 to 254 nests per year). Yellowhammers declined substantially in the UK after about 1985, a decline that may partly reflect reduced breeding frequency (Siriwardena *et al.* 2000). To test this hypothesis, we fitted a mixture model to yearly distributions of observed breeding dates, estimated $p_{1\text{est}}$ and hence B_{est} for each year to test whether B_{est} declined after 1985.

RESULTS

Validation of initial assumptions using individual-based data

We isolated the first and the subsequent breeding attempts from each intensive study (Mandarte song sparrows and Scottish yellowhammers) and tested whether a normal or a skew-normal model fitted each data subset best. Across the long-term song sparrow data, laying dates for first breeding attempts only were best fitted by a skew-normal distribution with location $\mu = -35.8$, scale $\sigma = 17.3$, shape $\alpha = 2.5$ (posterior means of MCMC fits) according to model DIC (Fig. 1a and see Table S3.1 in Appendix S3). Dates for all the subsequent breeding attempts (excluding data corresponding to first attempts) were fitted best by a normal distribution with $\mu = 15.5$ and $\sigma = 17.5$ (Fig. 1a). Across the intensive Scottish yellowhammer study, the distribution of laying dates was fitted best by a skew-normal distribution with $\mu = 35.4$, $\sigma = 17.9$, $\alpha = 4.7$, and a normal distribution with $\mu = 94.9$, $\sigma = 17.0$, for first and subsequent attempts respectively (see Fig. 1b and Table S3.1 in Appendix S3). Assumptions (2) and (3) were therefore supported for both species. Inspection of single years' data clearly indicates bimodality in every data set, and we therefore do not consider a unimodal model.

Validation of mixture model using song sparrow data

Across the long-term song sparrow data, the grand mean number of breeding attempts made per female per year (B_{obs}) was 2.51 ± 0.29 (mean \pm SD, ranging from 2.15 to 3.50). The equivalent summaries produced by the mixture model, namely the mean and standard deviation of annual B estimates' posterior mean, were 2.50 and 0.18 respectively (ranging from 2.16 to 2.86). Hence, the estimated grand mean B was indistinguishable from the observed grand mean B_{obs} .

Comparison of B_{est} and B_{obs} showed that annual estimates were relatively imprecise but correlated with the corresponding observed values ($r = 0.32$). Five points were poorly estimated, of which two were major outliers (Fig. 2a,b). Poor or aberrant estimates were easily spotted by examining model fit to laying date distributions in individual years (see examples in Fig. 2c–f). Poor fit was typically due to laying dates not being distinctly bimodal (first and successive attempts became confounded hence violating initial assumption (1) in years with high numbers of nests (e.g. 1985 and 1986, Fig. 2d) or to very atypical breeding phenology (e.g. 1990, Fig. 2e).

Effect of sample size on estimates of breeding frequency

We simulated data sets with total sample size N varying from 400 to as few as 25 nest records, in order to evaluate

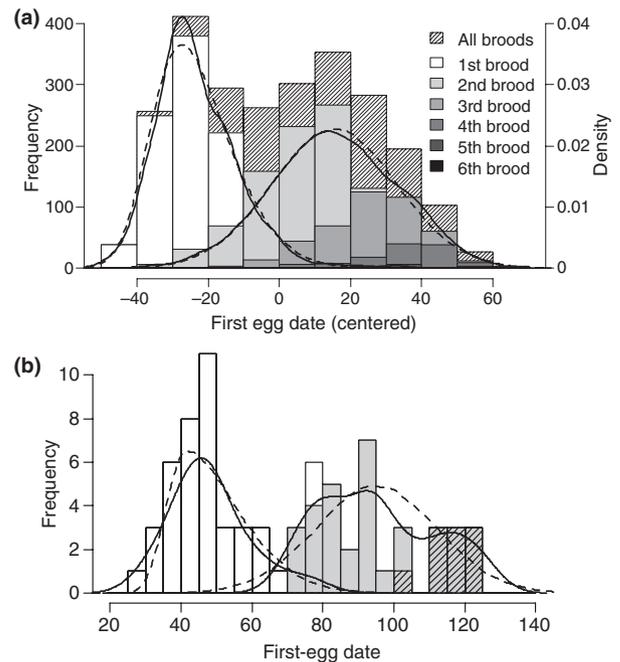


Figure 1 (a) Distribution of song sparrow first egg dates on Mandarte Island from 1975 to 2005. First egg dates have been centred within each year in order to remove annual variation in the timing of breeding. Successive broods are superimposed (higher order broods front most). Continuous lines represent kernel density estimates of the data for first broods (left) and subsequent broods (right). Dashed lines are modelled densities: left, assuming a skew-normal distribution with location $\mu = -35.8$, standard deviation $\sigma = 17.3$, shape $\alpha = 2.5$; and right, assuming a normal distribution with $\mu = 15.5$ and $\sigma = 17.5$. (b) Distribution of laying dates for first and subsequent breeding attempts for yellowhammers in the Scottish study sites (Day 1 = 1 April). White: first attempts. Grey: subsequent attempts. Shading lines show third attempts. Continuous and dashed lines are kernel smoothers of empirical and modelled distributions respectively. Laying dates were modelled assuming a skew normal distribution with $\mu = 35.4$, $\sigma = 17.9$, $\alpha = 4.7$, and a normal distribution with $\mu = 94.9$, $\sigma = 17.0$, for first and subsequent attempts respectively (see Results).

the performance of our method with small data sets. In our simulations, the estimates of the number of reproductive attempts tend to be slightly biased low (p_1 overestimated) for all sample sizes (Fig. 3). For sample sizes at least 100, the coefficient of variation is near 0.11 and precision improves only marginally by increasing sample size.

Application of mixture model to yellowhammer data

Application of the mixture model to the long-term UK nest record scheme data produced estimates of B from 1960 to 2004. These models suggested that nesting phenology differed markedly before the start of population decline

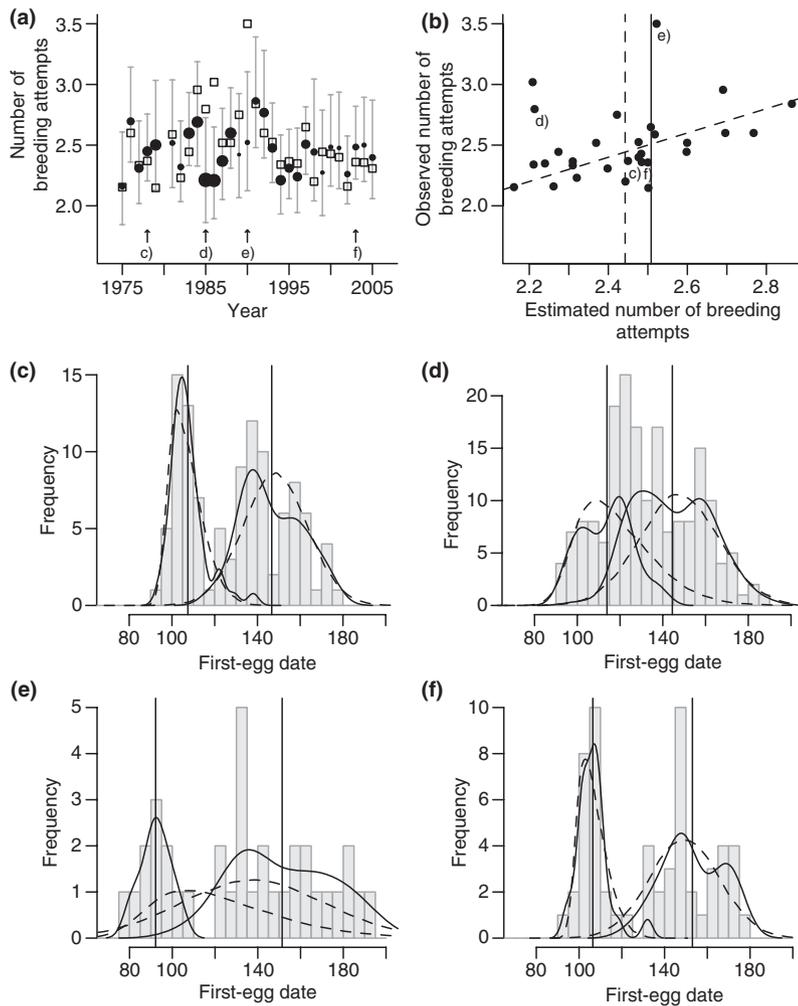


Figure 2 Model output for Mandarte song sparrows, comparing (a) observed (crosses in squares) and estimated (dots, with 95% credible interval) number of breeding attempts per year over time. Dot size reflects the total number of nests found. Arrows point to the years detailed in panels (c)–(f); (b) Relationship between observed and estimated number of breeding attempts per year. The diagonal shows equality between estimates and observed values, vertical lines show 1975–2005 averages for observed (thin continuous) and estimated values (thick dashed); (c–f) Typical examples of first egg date distributions for individual years: smooth lines represent density estimates (observed and fitted by the model) as in Fig. 1, vertical lines, left to right, are mean first egg date for observed first and subsequent broods respectively. (c) and (f) are typical years with satisfactory model fit; (d) is an example of poor model fit due to a unimodal distribution of first egg dates; (e) shows an atypical distribution of first egg dates, with exceptionally early first broods, classical dates for re-nesting and very few nests in total.

compared to after the decline was occurring (around 1985), with a smaller proportion of nests being first attempts after 1985 (Fig. 4). The grand mean B was 1.66 ± 0.27 , with estimates ranging from 1.27 to 2.37 across years. Although no trend over time was formally included in the model of breeding frequency, the results certainly do not support the hypothesis that the number of breeding attempts is declining: on the contrary, our model provides more support for an increase in the number of breeding attempts concurrent with the period of yellowhammer population decline (Figs 4 and 5).

DISCUSSION

Census data from population monitoring programmes often allow estimation of demographic parameters including survival and reproductive success (brood or litter size). However, estimating the number of breeding attempts for the many taxa that can breed more than once per season (including insects, fish, small birds and mammals) is

problematic. Here we proposed a new Bayesian mixture-model method to disentangle the overlapping distributions. We then tested this on simulated data and data from two bird species. Application of the method to a large data set of a farmland bird in decline indicates that the decline is not driven by a reduction in reproductive performance.

Successes and limitations of the approach

Our mixture model successfully estimated the grand mean number of breeding attempts made per individual per year (B_{obs}) in the intensively studied song sparrow population on Mandarte Island. Annual estimates of B were moderately correlated with observed B_{obs} , which is notable given the limited data available for some years from this small population. Whether model parameters are best considered as random or fixed effects is a modelling decision that will be affected partly by what is believed about variation between years, partly by the amount of data available for each year, as sharing the available information between years

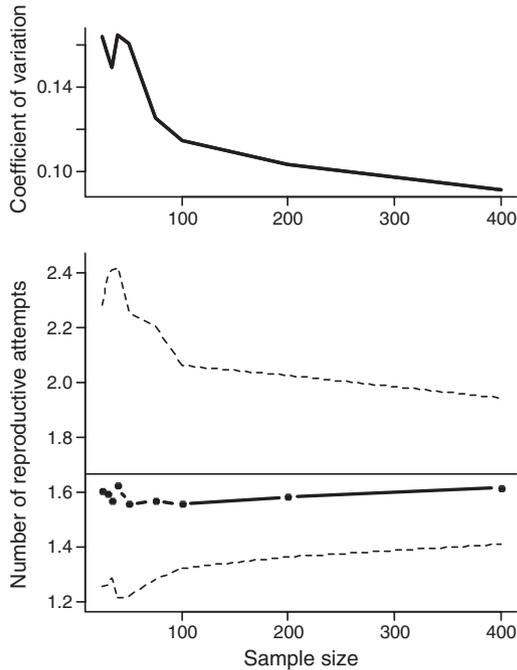


Figure 3 Effect of sample size on estimated number of reproductive attempts. Upper panel: coefficient of variation of the posterior mean number of reproductive attempts over 100 simulations, against sample size. Lower panel: median (thick line), 2.5% and 97.5% quantiles (dashed lines) of the 100 posterior means for the number of reproductive attempts ($1/p_1$). The horizontal thin line shows the true value used in the simulations.

will produce more robust estimates, and partly by the quality of fit to the data. Assuming that year-specific values of p_1 , say p_{1i} , are drawn from the same hyperdistribution over all years comes with both the risks and benefits of shrinking the estimates of p_{1i} towards the mean. This is best exemplified by the close to average song sparrow estimate of B in 1990, while the true number of nesting attempts was exceptionally high, but the amount of information was very limited due to low population size. Other poor estimates of B between the mid-1980s and 1990s are related to years with high overlap between the distribution of first and subsequent breeding attempts that reduced our ability to

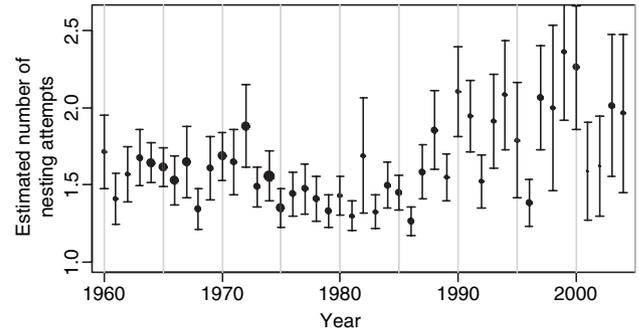


Figure 5 Modelled number of breeding attempts per year in the yellowhammer with 95% credible interval. Dot size reflects the relative number of nest record cards returned for each year.

distinguish among them and, thus, our capacity to estimate B accurately. Here it was our choice to keep the model very simple in order to demonstrate the potential and limits of the approach. In studies where the precision of individual estimates is essential and the amount of information sufficient, further effort may be put on accommodating the idiosyncrasies of the data set, by allowing adequate parameters to vary. We have in fact tried different combinations of fixed and random effects with the Mandarte's song sparrow data set (Table S3.3 in Appendix S3). We found that treating all year-specific parameter values as fixed effects except one (p_{1i}) seemed to work best for us, as otherwise some of the corresponding estimates of B_i were highly unrealistic. However, there is no reason for this to be a general rule and some methodological work is required following our demonstration of the potential for this approach.

This suggests that our method, within comparable conditions (i.e. number of data available and identifiability of the underlying distributions), is at least suitable for detecting broad trends in the number of breeding attempts, but that individual annual estimates should be treated more cautiously when sample size is small or the assumptions of the method are violated. Despite its large total size, the Mandarte data set is a relatively difficult example because the number of nests per year is limited (25% of years yield

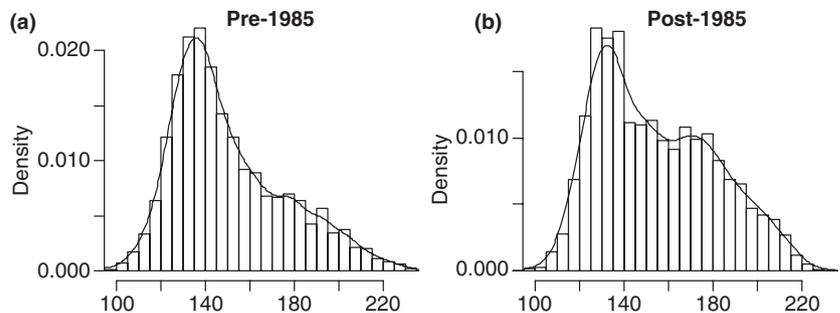


Figure 4 Distribution of yellowhammer first egg dates in the UK, with kernel density estimate. (a) From 1960 to 1985, (b) from 1986 to 2004.

fewer than 55 nests, while only 25% exceed 115 nests), and in some years the distribution of laying dates shows little evidence of bimodality. However, we find using simulations that our method gives reasonable estimates even with relatively small sample sizes (i.e. 100 data) when model assumptions are satisfied. A comparison of fitted against empirical distributions of laying dates is essential in order to assess the reliability of B estimates. Many factors may influence the precision of the results, either alone or in combination, including sample size, the adequacy of the hypothetical underlying distributions, the separation of the two distributions and seasonal biases in reporting. Note also that the estimation errors tend to increase for smaller values of p_1 , that is for larger numbers of breeding attempts B (see Fig. S2.2 in Appendix S2). Both the increase in B and a reduction in the number of nest records may explain the larger variance in the yellowhammer estimates after 1985 in the UK.

The ability to distinguish first breeding attempts from subsequent ones depends on the combination of two main factors: (1) the difference in the mean breeding date between these two groups; (2) the variance about the mean for each group (the degree to which individuals breed in synchrony). Distinguishing first from second will become easier as (1) increases and (2) decreases. Therefore, species with short incubation + rearing + inter-attempt times will thus tend to be more difficult to estimate unless the variance in breeding dates is small. Long inter-attempt times should always help discriminating first and subsequent attempts. On the other hand, relatively long incubation + rearing times should amplify the effect of failure rate by inducing larger differences in re-breeding dates between individuals that failed early and those that did not.

Is a reduction in breeding attempts the cause of the UK yellowhammer decline?

The utility of our method is illustrated by the insight gained into the reasons for farmland bird declines. Bradbury *et al.* (2000) suggested that yellowhammers undertake up to three breeding attempts, with a maximum of two successful ones. Our large-scale estimates from the BTO data indicate on average 1.66 breeding attempts per year for the yellowhammer, ranging from 1.27 to 2.37. This average may appear low given that it includes failed attempts. In addition, late nests are known to be under-represented in the BTO's Nest Record Scheme, with volunteer search effort tailing off after July (Crick *et al.* 2003), whereas the yellowhammer breeding season extends into August. As a consequence, late nests are possibly under-represented and B is possibly underestimated. However, our estimates at the end of the period are consistent with our recent field data from Scotland (2005–2006) that indicate an average of 2.31 attempts per pair, of

which 1.62 were successful. Ten per cent of pairs achieved three successful attempts (see Table 1).

Our results show evidence for an increase in the relative number of breeding attempts after 1985 that is concurrent with the yellowhammer population decline in the UK. This trend is robust to seasonal bias in nest records provided that the bias did not change consistently over time, and there is no evidence that this has occurred (R.A. Robinson, unpublished data).

This finding has two major consequences. First, a reduction in breeding frequency cannot be invoked to explain the decline of the yellowhammer, as suggested previously (Bradbury *et al.* 2000; Siriwardena *et al.* 2000). In the decade before the population decline (1975–1985) there averaged ~ 1.3 breeding attempts per season, in the decade after there were ~ 2.3 . This increase in breeding frequency was concomitant with a twofold increase in yellowhammer productivity per nesting attempt, due to a nearly 50% decline in nest failure rate that preceded population decline and which has remained at its lowest level since then (Siriwardena *et al.* 2000). Therefore, total productivity per nesting pair in the yellowhammer should have more than doubled as its population has declined. Second, this increase in both productivity per nesting attempt and breeding frequency suggests a decrease in competition for breeding resources between territorial pairs. In terms of yellowhammer conservation, the current high productivity per nesting pair suggests that efforts to improve non-breeding habitat quality may be more relevant at a large scale than those directed at increasing breeding output (see for instance, Siriwardena *et al.* 2007). Nevertheless, breeding habitat improvement (see for instance, Douglas *et al.* 2009) may be useful in areas where population density remains high and competition for resources is likely.

Future directions

We aimed to develop and assess a novel method for estimating breeding frequency, and use it to address long-standing

Table 1 Distribution of number of nesting attempts per breeding pair of yellowhammer in the Scottish study sites (restricted to pairs with all attempts known, $N = 29$). Figures are given in percentage of pairs. Average number of attempts per pair was 2.31 and average number of successful attempts per pair was 1.62

Number of attempts	Number of successful attempts				Total
	0	1	2	3	
1	3.4	3.4	0.0	0.0	6.9
2	3.4	10.3	41.4	0.0	55.2
3	3.4	13.8	10.3	10.3	37.9
Total	10.3	27.6	51.7	10.3	–

questions about limits on populations of farmland birds (see for instance, Newton 2004). In the case of UK birds, more work is needed to estimate better the absolute number of breeding attempts, in ways that correct for seasonal observation bias and parameterize demographic models. This next step might be achieved by taking fuller advantage of Bayesian approaches to combine Nest Record data from multiple species with information from a smaller subset of species that have been the subject of intensive studies. Because local, intensive studies should have little or no bias in sampling effort, they could be used to correct the seasonal sampling bias in NRS volunteer-based records across species.

In a more general context, a large-scale simulation-based assessment of the approach would be useful, for instance, to assess quantitatively the consequences of departures from simple initial assumptions, or the effects of sampling bias. An obvious extension would involve including relevant covariates on breeding date, for example altitude, latitude, weather, or the presence or abundance of major brood predators (Brown *et al.* 1999; Crick & Sparks 1999; Sanz 2003; Gordo *et al.* 2007). Relationships between breeding phenology and nest failure rates are not straightforward ecologically (Ratcliffe *et al.* 2005; Weggler 2006) or statistically. Nevertheless, including information on nest failure rates, for which these data sets are often collected, may help estimating the mean and standard deviation of laying dates, particularly for late clutches.

CONCLUSION

A merit of the approach was its ability to produce precise estimates in spite of the relatively heterogeneous data and the severe violation of its assumptions in several years. This was indeed achieved by pooling information among years through the use of random effects, while simulation suggests that reasonable precision may be achieved with about 75 nest records in single-year data sets. The method proved useful in addressing the role of variation in the number of breeding attempts in a declining farmland bird, the yellowhammer. This method is applicable to a number of multi-brooded bird species, to single brooded species with replacement clutches and to non-avian species with repeated reproductive events per season. In particular, however, our method should allow researchers to revisit and use more thoroughly historical data sets to extract key demographic parameters essential for predicting the temporal and spatial dynamics of populations.

ACKNOWLEDGEMENTS

This work is the result of a NERC grant awarded through UKPopNet to W.J.S., T.G.B. and X.L. William J. Sutherland is funded through the Arcadia Fund. Authors other than T.C.

and D.A.E. are in alphabetical order rather than in order of contribution. We thank all the volunteers who contributed to the Nest Record Scheme with the British Trust for Ornithology. The NRS is funded by a partnership between the BTO and the Joint Nature Conservation Committee. We thank NSERC and NSF for funding the song sparrow study, and the Tsawout and Tseycom first nations bands for permission for that study to work on Mandarte.

REFERENCES

- Aebischer, N.J. (1999). Multi-way comparisons and generalised linear models of nest success: extensions of the Mayfield method. *Bird Study*, 46(Suppl.), S22–S31.
- Bradbury, R.B., Kyrkos, A., Morris, A.J., Clark, S.C., Perkins, A.J. & Wilson, J.D. (2000). Habitat associations and breeding success of yellowhammers on lowland farmland. *J. Appl. Ecol.*, 37, 789–805.
- Brown, J.L., Li, S.-H. & Bhagabati, N. (1999). Long-term trend toward earlier breeding in an American bird: a response to global warming? *Proc. Natl. Acad. Sci. U.S.A.*, 96, 5565–5569.
- Browne, S.J. & Aebischer, N.J. (2004). Temporal changes in the breeding ecology of European turtle doves *Streptopelia turtur* in Britain, and implications for conservation. *Ibis*, 146, 125–137.
- Congdon, P. (2001). *Bayesian Statistical Modelling*. Wiley, Chichester.
- Crick, H.Q.P. & Sparks, T.H. (1999). Climate change related to egg-laying trends. *Nature*, 399, 423–424.
- Crick, H.Q.P., Baillie, S.R. & Leech, D.I. (2003). The UK Nest Record Scheme: its value for science and conservation: a review of its methodology, analytical procedures and uses. *Bird Study*, 50, 254–270.
- Donald, P.F., Green, R.E. & Heath, M.F. (2001). Agricultural intensification and the collapse of Europe's farmland bird populations. *Proc. R. Soc. Biol. Sci. Ser. B*, 268, 25–29.
- Donald, P.F., Evans, A.D., Muirhead, L.B., Buckingham, D.L., Kirby, W.B. & Schmitt, S.I.A. (2002). Survival rates, causes of failure and productivity of skylark *Alauda arvensis* nests on lowland farmland. *Ibis*, 144, 652–664.
- Douglas, D.J.T., Vickery, J.A. & Benton, T.G. (2009). Improving the value of field margins as foraging habitat for farmland birds. *J. Appl. Ecol.*, 46, 353–362.
- Gordo, O., Sanz, J. & Lobo, J. (2007). Spatial patterns of white stork (*Ciconia ciconia*) migratory phenology in the Iberian Peninsula. *J. Ornithol.*, 148, 293–308.
- Hansson, B., Bensch, S. & Hasselquist, D. (2000). The quality and the timing hypotheses evaluated using data on great reed warblers. *Oikos*, 90, 575–581.
- Johnson, D.H. (1979). Estimating nest success: the Mayfield method and an alternative. *Auk*, 96, 651–661.
- Lande, R., Engen, S. & Saether, B.E. (2003). *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, Oxford.
- Lebreton, J.-D., Burnham, K.P., Clobert, J. & Anderson, D.R. (1992). Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecol. Monogr.*, 62, 67–118.
- Mayfield, H. (1961). Nesting success calculated from exposure. *Wilson Bull.*, 73, 255–261.

- Nagy, L.R. & Holmes, R.T. (2005). Food limits annual fecundity of a migratory songbird: an experimental study. *Ecology*, 86, 675–681.
- Newton, I. (2004). The recent declines of farmland bird populations in Britain: an appraisal of causal factors and conservation actions. *Ibis*, 146, 579–600.
- Nichols, J.D., Kendall, W.L. & Runge, M.C. (2005). Estimating survival and movement. In: *Bird Ecology and Conservation* (eds Sutherland, W.J., Newton, I., Green, R.E.). Oxford University Press, Oxford, pp. 119–140.
- Pease, C.M. & Grzybowski, J.A. (1995). Assessing the consequences of brood parasitism and nest predation on seasonal fecundity in passerine birds. *Auk*, 112, 343–363.
- R Development Core Team (2007). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ratcliffe, N., Schmitt, S. & Whiffin, M. (2005). Sink or swim? Viability of a black-tailed godwit population in relation to flooding. *J. Appl. Ecol.*, 42, 834–843.
- Sanz, J.J. (2003). Large-scale effect of climate change on breeding parameters of pied flycatchers in Western Europe. *Ecography*, 26, 45–50.
- Siriwardena, G.M., Baillie, S.R., Crick, H.Q.P. & Wilson, J.D. (2000). The importance of variation in the breeding performance of seed-eating birds in determining their population trends on farmland. *J. Appl. Ecol.*, 37, 128–148.
- Siriwardena, G.M., Stevens, D.K., Anderson, G.Q.A., Vickery, J.A., Calbrade, N.A. & Dodd, S. (2007). The effect of supplementary winter seed food on breeding populations of farmland birds: evidence from two large-scale experiments. *J. Appl. Ecol.*, 44, 920–932.
- Smith, J.N.M. (2006). Song sparrows, Mandarte Island, and methods. In: *Conservation and Biology of Small Populations: The Song Sparrows of Mandarte Island* (eds Smith, J.N.M., Keller, L.F., Marr, A.B., Arcese, P.). Oxford University Press, New York, pp. 18–29.
- Thomas, A., O'Hara, B., Ligges, U. & Sturtz, S. (2006). Making BUGS open. *R News*, 6, 12–17.
- Thompson, B.C., Knadle, G., Brubaker, D.L. & Brubaker, K.S. (2001). Nest success is not an adequate comparative estimate of avian reproduction. *J. Field Ornithol.*, 72, 527–536.
- Thomson, D.L. & Cotton, P.A. (1999). *Understanding the Decline of the British Population of Song Thrushes Turdus philomelos*. British Ornithologists' Union, University of Southampton, UK, pp. 151–155.
- Weggler, M. (2006). Constraints on, and determinants of, the annual number of breeding attempts in the multi-brooded Black Redstart *Phoenicurus ochruros*. *Ibis*, 148, 273–284.
- Wilson, S. & Arcese, P. (2003). El Niño drives timing of breeding but not population growth in the song sparrow (*Melospiza melodia*). *Proc. Natl. Acad. Sci. U S A*, 100, 11139–11142.

SUPPORTING INFORMATION

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

Appendix S1 BUGS code for the mixture model.

Appendix S2 Simulated data and mean–variance relationship of the estimate.

Appendix S3 Model comparisons.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

Editor, Jordi Bascompte

Manuscript received 17 April 2009

First decision made 25 May 2009

Manuscript accepted 24 July 2009