



Analysis of Population Trends for Farmland Birds Using Generalized Additive Models

Rachel M. Fewster; Stephen T. Buckland; Gavin M. Siriwardena; Stephen R. Baillie;
Jeremy D. Wilson

Ecology, Vol. 81, No. 7 (Jul., 2000), 1970-1984.

Stable URL:

<http://links.jstor.org/sici?sici=0012-9658%28200007%2981%3A7%3C1970%3AAOPTFF%3E2.0.CO%3B2-4>

Ecology is currently published by The Ecological Society of America.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/esa.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

ANALYSIS OF POPULATION TRENDS FOR FARMLAND BIRDS USING GENERALIZED ADDITIVE MODELS

RACHEL M. FEWSTER,^{1,4} STEPHEN T. BUCKLAND,¹ GAVIN M. SIRIWARDENA,^{2,3} STEPHEN R. BAILLIE,²
AND JEREMY D. WILSON^{3,5}

¹*Department of Mathematical and Computational Sciences, University of St Andrews, Mathematical Institute,
North Haugh, St Andrews, Fife KY16 9SS, UK*

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

³*Ecology and Behaviour Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK*

Abstract. Knowledge of the direction, magnitude, and timing of changes in bird population abundance is essential to enable species of priority conservation concern to be identified, and reasons for the population changes to be understood. We give a brief review of previous techniques for the analysis of large-scale survey data and present a new approach based on generalized additive models (GAMs). GAMs are used to model trend as a smooth, nonlinear function of time, and they provide a framework for testing the statistical significance of changes in abundance. In addition, the second derivatives of the modeled trend curve may be used to identify key years in which the direction of the population trajectory was seen to change significantly. The inclusion of covariates into models for population abundance is also discussed and illustrated, and tests for the significance of covariate terms are given. We apply the methods to data from the Common Birds Census of the British Trust for Ornithology for 13 species of farmland birds. Seven of the species are shown to have experienced statistically significant declines since the mid-1960s. Two species exhibited a significant increase. The population trajectories of all but three species turned downward in the 1970s, although in most cases the 1980s brought either some recovery or a decrease in the rate of decline. The majority of populations have remained relatively stable in the 1990s. The results are comparable with those from other analysis techniques, although the new approach is shown to have advantages in generality and precision. We suggest extensions of the methods and make recommendations for the design of future surveys.

Key words: *bird census schemes; British birds; change points; Common Birds Census; farmland birds; generalized additive models; log-linear Poisson regression; nonlinear trend; population trajectory; spatiotemporal models; trend analysis.*

INTRODUCTION

Accurate analysis of long-term monitoring data is essential for the effective management and conservation of wildlife populations. An important component of such analysis is the determination of trend in population abundance over time. However, the nature of ecological data is such that obtaining reliable estimates of annual abundance is far from straightforward. The methods that are usually employed for time series data are generally inappropriate, because the time span of the surveys is too short (Bowerman and O'Connell 1987). Difficulties are also caused by missing data, which tend to be characteristic of large-scale census schemes (ter Braak et al. 1994, Thomas 1996). Barker and Sauer (1992) outline some other problems typically associated with wildlife survey data.

In the past, indices of abundance were obtained from

census data using the chain method (e.g., Marchant et al. 1990). The ratio of abundance in one year to abundance in the next was calculated as the ratio of summed counts in the two years, with the summation taken over only those sites that were surveyed in both years. This method is now generally regarded as inadequate, because of inefficient use of the data and a tendency to generate spurious trends through random drift (Mountford 1982, 1985, Peach and Baillie 1994, ter Braak et al. 1994). A number of alternative procedures have since been proposed, particularly for bird census schemes (ter Braak et al. 1994, Thomas 1996). Nonetheless, there is still some scope for improvement (Thomas 1996).

In North America, most analyses are conducted using route regression techniques. For each of many surveyed routes, a linear regression of log-count against time is performed to yield a log-linear, route-specific trend. Overall trends, for example over a region or state, are obtained by combining the route-specific results using various weighting schemes (Geissler and Noon 1981, Geissler and Sauer 1990, Sauer and Geissler 1990). In Europe, attention has focused on *sites-by-years* models, in which the expected count in a given site and year

Manuscript received 2 June 1998; revised 14 May 1999; accepted 4 June 1999; final version received 28 June 1999.

⁴ Present address: Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand. E-mail: r.fewster@auckland.ac.nz

⁵ Present address: Royal Society for the Protection of Birds, The Lodge, Sandy, Bedfordshire SG19 2DL, UK.

is modeled as a function of quantities known as the *site effect* and the *year effect*. The site effect allows for variation in abundance between plots, whereas the year effect allows for fluctuations in abundance over time. These models have been fitted using Poisson regression techniques (Pannekoek and van Strien 1996) and by the Mountford method (Mountford 1982, 1985). In contrast to the log-linear form prescribed by route regression, the estimates of annual abundance derived from sites-by-years models are not constrained to follow any pre-specified curve or shape.

The distinct approaches largely reflect differences in the objectives of the analyses. Models that treat log-counts as a linear function of time are designed to estimate some average rate of population change over a time period, although the fit is likely to be poor if the true pattern of change was markedly nonlinear. Conversely, unconstrained annual abundance estimates reveal every fluctuation in population numbers, but can be hard to interpret in terms of long-term change. Consequently, neither approach is well-suited to the investigation of long-term, but nonlinear, trends in population numbers.

Previous attempts to distinguish genuine patterns of nonlinear change from the "noise" of annual fluctuation have centered on the application of smoothing algorithms to annual abundance estimates. For example, Siriwardena et al. (1998a) used a compound running-median algorithm for the smoothing of Mountford indices derived from British bird census data. However, smoothing the output from one model amounts to the application of a second model, and methods that incorporate nonlinear trend estimation directly into the fitting of the original model would be preferable.

In this paper, we present new methods for the analysis of trends in wildlife census data, using generalized additive models (GAMs). GAMs are not simply smoothing devices, but provide a complete modeling framework. Smoothing procedures are built into the model-fitting process, so that inference based on the resulting smooth curve of abundance indices is made fully within the context of the original model. Instead of requiring a linear form for the log trend, GAMs allow any shape ranging from a straight line, through a range of nonparametric curves of increasing complexity, to unconstrained annual estimates. These latter are equivalent to the estimates obtained for sites-by-years models from Poisson regression. The range of smooth curves available from the GAM allows analyses of counts to be tailored a priori according to the biological question of interest: simple linear changes, long-term nonlinear trends, or annual fluctuation.

Smoothing approaches have also been developed in North America to overcome the restrictions of linearity in the traditional route regression method. For example, Taub (1990) and James et al. (1990, 1996) have applied the smoothing algorithm LOESS to counts at the site

level and higher. This method has some similarities with the GAM approach, and these will be discussed.

We begin the present paper by formulating the GAM and demonstrating how the conventional sites-by-years model may be regarded as a special case. Significance tests are described for drawing inference from the smooth series of GAM abundance estimates. Next, we present a method to identify key years in which there was seen to be a significant change in the direction of the population trajectory. This is accomplished by estimating the curve of second derivatives of the abundance index curve, and identifying years in which the second derivative was significantly different from zero. Estimating the timing of population changes in this way helps to suggest causes of the change. We conclude with a discussion of covariate models, which enable factors such as geographical location or climate to be taken into account.

The methods are illustrated using data from the British Trust for Ornithology's Common Birds Census (CBC), although they are applicable to a wide range of census schemes. The CBC is one of the longest running of all wildlife monitoring schemes, with annual data available since 1962. Plots of land, or sites, are selected by volunteer observers for repeated visits throughout the breeding season. Territories of all bird species are mapped as accurately as possible according to standard criteria, and the final site count for a given species is the estimated number of territory-holding males in the site. The survey design is described in detail by Marchant et al. (1990).

GAMs are used in this paper to analyze CBC data from farmland plots for 13 British bird species, all predominantly granivorous. These are the Grey Partridge *Perdix perdix* L., Stock Dove *Columba oenas* L., Turtle Dove *Streptopelia turtur* L., Skylark *Alauda arvensis* L., Chaffinch *Fringilla coelebs* L., Corn Bunting *Miliaria calandra* L., Goldfinch *Carduelis carduelis* L., Greenfinch *C. chloris* L., Linnet *C. cannabina* L., Bullfinch *Pyrrhula pyrrhula* L., Tree Sparrow *Passer montanus* L., Reed Bunting *Emberiza schoeniclus* L., and Yellowhammer *E. citrinella* L. There is much concern about the status of many of these species, following declines across Europe in both abundance (Marchant and Gregory 1994, Tucker and Heath 1994, Fuller et al. 1995, Siriwardena et al. 1998a) and range (Gibbons et al. 1993, Tucker and Heath 1994, Fuller et al. 1995). In particular, population numbers are thought to have been heavily influenced by changes in farming practice since the 1960s. A primary objective is to identify the timing and extent of population declines, so that these can be linked to possible causal factors. For this reason, the fitting of a simple linear trend model would not be adequate, and the analysis demands the greater flexibility offered by the GAM approach.

METHODS

Formulation of a log-linear poisson regression model for CBC counts

The application of log-linear Poisson regression techniques to the analysis of bird census data is outlined by ter Braak et al. (1994), who considered the approach to be the most promising of those available at the time. The specialized free software package TRIM (Trends and Indices for Monitoring Data; Pannekoek and van Strien 1996) provides an efficient implementation of log-linear Poisson regression for data from census schemes, although most of the features of TRIM are also available in the generalized linear model modules of well-known packages such as S-PLUS (Statistical Sciences 1993), GLIM (Francis et al. 1993), or SAS (SAS Institute 1996). We show here how the log-linear Poisson model is formulated. In the next section, the formulation will be extended to yield a system of generalized additive models.

Data from the Common Birds Census may be regarded as a set of N sites monitored over T years. The number of sites N differs between species, because only those plots on which the species was observed are included in the model. For each species, the notation y_{it} is used to denote the number of territories, or *count*, observed in site i in year t .

In general, each CBC site was selected and surveyed by only one observer; the length of site coverage was therefore highly variable and ranged from two years to 33 years for different sites. The mean length of site coverage was 6–8 years for each of the 13 species' data sets. The remainder of site-by-year records are treated as missing data and must be estimated through the model. On the rare occasions when a site was surveyed by more than one observer over the time period, the results from the different observers are treated as records from separate sites. This ensures that effects due to differences in observer ability are not interpreted as genuine population changes.

The log-linear Poisson model is formulated by assuming that the observation y_{it} derives from a Poisson distribution with mean μ_{it} , and that all counts are independent. The mean μ_{it} is modeled as follows:

$$\log(\mu_{it}) = \alpha_i + \beta_t \quad (1)$$

The quantities α_i and β_t are referred to as the site effect for site i and the year effect for year t , respectively.

The one-to-one relationship between observers and sites in the CBC means that observer effects are indistinguishable from site effects in the model given by Eq. 1. In effect, a parameter is included for every pairing of site and observer. With some other surveys, a slightly different formulation might be appropriate. If it were common not only for each site to be surveyed by several different observers, but also for each observer to survey many different sites, it might be possible to estimate observer effects separately from site

effects by including a parameter for each in Eq. 1. This would also allow identification of observers who consistently recorded particularly high or low counts.

The model specified by Eq. 1 belongs to a class known as generalized linear models (GLMs), and has Poisson error distribution and logarithmic link function. GLMs are described in detail by McCullagh and Nelder (1989). The model is fitted by finding estimates $\hat{\alpha}_i$ and $\hat{\beta}_t$ for the N parameters α_i and the T parameters β_t . Once these estimates are obtained, they are used to provide an estimate of the mean count μ_{it} for each site i and year t , which can be interpreted as a prediction of the true count. The predicted count for site i in year t is

$$\hat{\mu}_{it} = \exp(\hat{\alpha}_i + \hat{\beta}_t) \quad (2)$$

and the total predicted count for year t is

$$\sum_{i=1}^N \hat{\mu}_{it} = \exp(\hat{\beta}_t) \sum_{i=1}^N \exp(\hat{\alpha}_i). \quad (3)$$

We define the *abundance index for year t* to be:

$$I_t = \frac{\text{total predicted count for year } t}{\text{total predicted count for year } 1} = \frac{\exp(\hat{\beta}_t)}{\exp(\hat{\beta}_1)} \quad (4)$$

The index measures relative abundance with respect to an arbitrarily chosen base year; in this case, year 1. Because the index is a ratio of exponentially transformed year effects, it is easily calculated without performing the sum in Eq. 3.

The index defined in Eq. 4 uses predicted counts for all sites, regardless of whether or not observations are available. With log-linear Poisson regression, the same index would be obtained if predictions were replaced by observations where possible, because the sum of the observations is equal to the sum of the associated predictions. This occurs because the logarithmic link function is canonical for the Poisson error distribution, and it is a general property of GLMs fitted with canonical link and including an intercept or factor (Nelder and Wedderburn 1972, ter Braak et al. 1994). A canonical link function exists for every permissible GLM error distribution (McCullagh and Nelder 1989).

From Eq. 4, it is clear that the index for the base year is always unity. This does not mean that there is no uncertainty associated with the abundance estimate in that year, but rather that the uncertainty is inherited by the rest of the index series. High uncertainty in the base year abundance estimate will induce wide confidence intervals throughout the series. If confidence intervals are unduly wide, an alternative choice for the base year might be considered.

The GLM formulated in Eq. 1 corresponds to the "annual model" of the TRIM software package (Pannekoek and van Strien 1996). The calculation of the annual abundance estimate for any year does not take into account the estimates for adjacent years, and in this sense the estimates are unconstrained. The model

lies at one extreme of the GAM framework, and is suitable for examining yearly fluctuations but not for discerning long-term trends.

Formulation of the generalized additive model

Model specification.—We now recast the log-linear Poisson GLM as a generalized additive model (GAM). GAMs are flexible extensions of GLMs, and are described comprehensively by Hastie and Tibshirani (1990). Once again, we assume that the counts y_{it} follow independent Poisson distributions, with mean μ_{it} for the count in site i in year t . However, the linear predictor associated with the GLM, which was given by Eq. 1, is now replaced by a more general *additive predictor*. This allows the change in mean abundance over time to follow any smooth curve, not just a linear form or a sequence of unrelated estimates as in Eq. 1. The form of the predictor function is the principal difference between the GLM and the GAM.

We write the additive predictor as

$$\log(\mu_{it}) = \alpha_i + s(t). \quad (5)$$

The expected count μ_{it} in site i in year t therefore depends upon the site effect α_i , and upon any number of other smoothly varying quantities, which are summarized by the value $s(t)$ in year t . The notation $s(t)$ simply represents any smooth function of time. The GAM is fitted by estimating the parameters α_i and the smooth function s , just as the GLM in Eq. 1 is fitted by estimating the parameters α_i and β_i .

There are two special cases of the GAM formulation that fall into the category of GLMs. The first is the simple linear trend model, in which $s(t) = \gamma t$ for a single parameter γ to be estimated. In this case, the expected abundance within each site varies linearly on a logarithmic scale with time. The second case is the log-linear Poisson regression model described in the previous section, for which $s(t) = \beta_i$ for parameters β_1, \dots, β_T to be estimated. In this instance, the function s is no longer smooth, and is obtained by joining the estimates β_i with straight lines.

These two cases lie at opposite extremes of the GAM framework. The first has maximum smoothness in the function s , a single straight line; the second has minimum smoothness, a sequence of unconstrained estimates joined by linear segments. Between the two extremes lie functions s with greater flexibility than the linear trend $s(t) = \gamma t$, but with smooth output in contrast to the discrete annual estimates $s(t) = \beta_i$. These intermediate GAM curves provide opportunities for eliciting long-term nonlinear trends that are not available at the GLM extremes.

The output from the GAM in Eq. 5 is visualized as follows. The fitted year effect curve $\hat{s}(t)$ is common to all sites, so that for any two sites i_1 and i_2 , the curves $\log(\mu_{i_1t})$ and $\log(\mu_{i_2t})$ are parallel. The intercepts of these curves are determined by the site effects, respectively α_{i_1} and α_{i_2} . Consequently, every site is subject to the

same trend in the logarithm of expected count over time, although the absolute values differ between sites.

The assumption that log abundance follows the same trend in every site, although simplistic, is the same as that implicit in the traditional Mountford and Poisson regression models. Indeed, the assumption is perhaps better justified in the GAM context, because it is more reasonable to suggest that there is a smooth pattern of change common to all sites than to assert that all sites are subject to the same series of annual fluctuations. Nonetheless, the GAM system also accommodates models that allow trends to vary between different regions, or even within every site, and these will be discussed in later sections.

Once an estimate \hat{s} has been obtained for the smooth function s in Eq. 5, the annual abundance index curve $I(t)$ is calculated as before:

$$I(t) = \frac{\text{total predicted count for year } t}{\text{total predicted count for year 1}} = \frac{\exp(\hat{s}(t))}{\exp(\hat{s}(1))}. \quad (6)$$

Note that I is now written as a smooth function of t , in contrast to the set of point estimates specified by Eq. 4.

Estimation of the smooth function s and specification of the degree of smoothing.—Before the function s can be estimated, the required level of smoothing must be established. At the two extremes of maximum and minimum smoothing, where the GAM is equivalent to a GLM, it is possible to specify s in parametric form: namely $s(t) = \gamma t$ at the upper extreme and $s(t) = \beta_i$ at the lower. Between these extremes, however, the function s is not defined in terms of parameters, but is estimated nonparametrically from the data by means of scatterplot smoothers. The shape of the function is therefore determined by the data rather than being restricted to a parametric form. Note the distinction between the smoothing procedures in a GAM, which are part of the model-fitting process, and the smoothing of Mountford or TRIM indices in a second step after the fit is complete.

Several options are available for the scatterplot smoother used in estimating s , although experiment suggests that the choice has little effect on the final index curve. All of the analyses presented in this paper were conducted using smoothing splines; other options include locally weighted regression smoothers, kernel smoothers, and running-median smoothers (Hastie and Tibshirani 1990).

Smoothing splines fit the data using piecewise cubic polynomials. They are chosen to satisfy a penalized least squares criterion, which optimizes the fit while penalizing roughness to some pre-specified extent. The extent to which roughness is penalized, or equivalently the level of smoothing that is applied, is calibrated by a quantity known as the *degrees of freedom* (df). As the degrees of freedom are increased, the function s gains in flexibility: more turning points and gradient

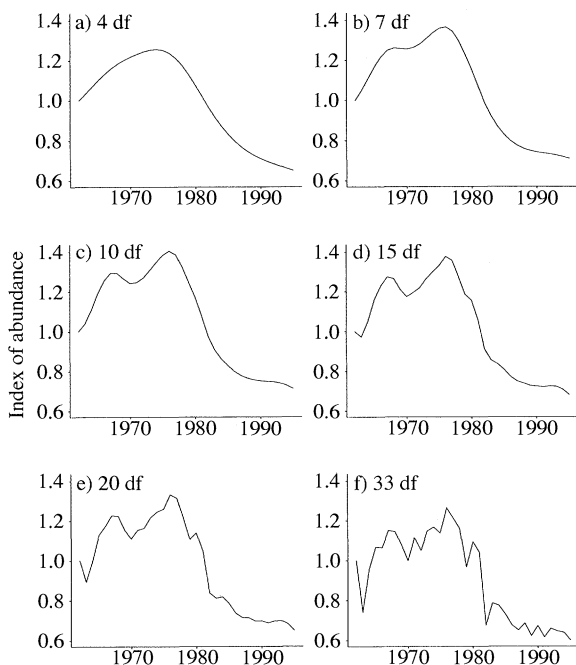


FIG. 1. Abundance indices from GAMs with various degrees of freedom fitted to CBC data for the Skylark.

changes are accommodated. A straight line $s(t) = \gamma t$ corresponds to a fit with 1 df; the curve follows a single direction only, with no gradient changes or turning points. Conversely, a piecewise linear fit $s(t) = \beta_i$, ($t = 1, \dots, T$) employs $T - 1$ df; a separate gradient is allowed between every successive pair of points. The degrees of freedom associated with the curve s may take any value between 1 and $T - 1$, and are loosely interpreted as the number of parameters used in fitting s .

The choice of the value for df is an important part of the modeling process. For clarity, we shall write s_d for the curve s to be fitted using a smoothing spline on the variable t with d degrees of freedom, and we shall refer to the associated model as a "GAM with d degrees of freedom." The choice of d depends largely on the objectives of the analysis. For inference about long-term trends, a smooth index curve is required, corresponding to low df; whereas information about annual fluctuations requires unconstrained annual estimates and the maximum value of d . Consideration of the length of the time series is also important, because longer time series will demand higher values of d if it is required to maintain a given level of flexibility in the trend curve.

With the CBC data, our objective was to detect all major features of population trend over the time period while ignoring fine-scale fluctuations. Fig. 1 shows how the abundance index curves obtained from the Skylark data change as the value of d is increased. Certain conceptual difficulties enter at this point, be-

cause it is not possible to assert that any one of these curves is more realistic than the others. The GAMs separate underlying trend from short-term fluctuation, but the precise point at which a signal can be considered fluctuation rather than trend is ill-defined. The different curves in Fig. 1 effectively represent models with different criteria for achieving this separation, and there is no right answer against which to judge them.

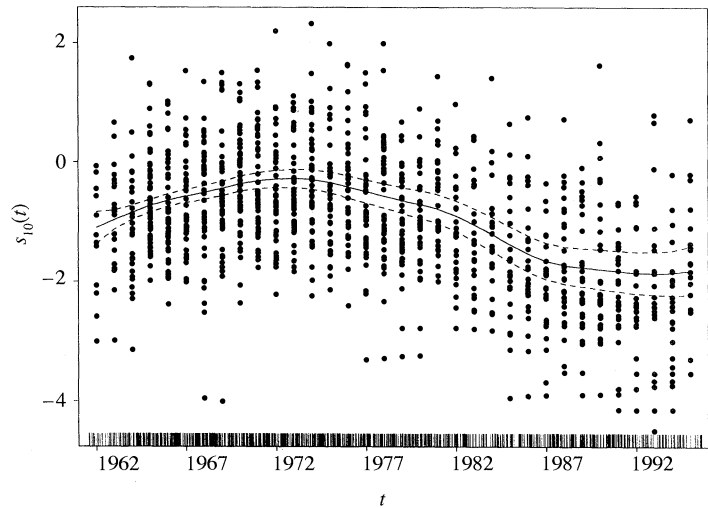
To overcome this difficulty, it is necessary to formulate our own ideas about a reasonable definition of noise and trend. As d is raised from the low starting value of 4 in Fig. 1, the index curve begins to reveal new features, but still remains smooth until a period of stability is reached around $d = 10$. Further increases in d serve only to roughen the output. It is reasonable to take those points at which the output has stabilized, but is still smooth, as our definition of true trend; our objectives were therefore fulfilled by a GAM with about 10 df. This value was also suitable for the other species examined in this paper. Experiments with truncation of CBC data suggested that a value for d of roughly 0.3 times the length of the time series tended to produce trend curves with suitable complexity and smoothness, although we stress that advice will vary according to precise objectives and data. It is always necessary to plot indices from GAMs with a range of df before settling on a final value.

Automatic selection of the value of df in a generalized additive model is sometimes recommended, using, for example, Akaike's Information Criterion (AIC) or generalized cross-validation (Hastie and Tibshirani 1990). Each automatic procedure produces a single value of d that is optimal with respect to some criterion. Because the quantity optimized might be very different from our own requirements of the model, however, we do not recommend the use of automatic selection procedures in the context of trend analysis. For example, we might wish to examine the same data both for long-term trends and for information about annual fluctuations. Although different values of d are clearly required for the two purposes, an automatic selection routine would be unable to distinguish between the distinct objectives, and would provide the same value for both.

Diagnostics

Once the GAM has been fitted, informal verification of the goodness-of-fit may be obtained through plots of residuals and standard errors, which are provided by most GAM software. Fig. 2 is an example of the diagnostic plot supplied by S-PLUS (Statistical Sciences 1993: version 3.4 for UNIX), which was the package used for the analyses in this paper. The solid line on the figure is the year effect curve $s_{10}(t)$ from a GAM with 10 df fitted to CBC data for the Corn Bunting, whereas the dashed lines lie two standard errors away from the central curve on either side. Calculation of the standard errors in this context involves a compli-

FIG. 2. Year effect curve $s_{10}(t)$ from a GAM with 10 df fitted to CBC data for the Corn Bunting (solid line), with 2 SE bands (dashed lines), partial deviance residuals (black dots), and rug plot. The residuals are well distributed above and below the curve; the standard error bands are mostly narrow, widening at the end points where the observations are more sparse.



cated series of approximations; details are given in Chambers and Hastie (1993).

The black dots on Fig. 2 represent *partial deviance residuals* (Chambers and Hastie 1993). The deviance is related to the difference in log-likelihood between the fitted model and a saturated model (McCullagh and Nelder 1989), and the deviance residual for observation y_{it} is the square root of the contribution made by y_{it} to the overall deviance. It is positive if y_{it} is greater than its fitted value $\hat{\mu}_{it}$, and negative if y_{it} is less than $\hat{\mu}_{it}$. With the Poisson error distribution used here, the partial deviance residual for observation y_{it} is given by

$$\text{sign}(y_{it} - \hat{\mu}_{it}) \sqrt{2 \left(y_{it} \log \frac{y_{it}}{\hat{\mu}_{it}} - (y_{it} - \hat{\mu}_{it}) \right)}. \quad (7)$$

Rather than displaying their absolute values, the diagnostic plot shows the residuals plotted about the fitted curve $s_{10}(t)$. The residuals farthest from the curve tend to represent observations with the lowest likelihoods under the fitted model. Within a Poisson framework, these will often be observations involving relatively low counts, because small fitted values imply a low Poisson variance and, consequently, less room for flexibility. The highest deviance residual in Fig. 2, for instance, occurs in the fit for year 1990, and it is traced to an observation $y_{it} = 3$ that has very low likelihood under its fitted value of 0.2.

A satisfactory diagnostic plot has residuals distributed evenly above and below the fitted curve, and standard error bands that are consistently narrow. Parts of the plot where the standard error bands are particularly wide, or where the residuals lie on only one side of the fitted curve or display some other obvious pattern, suggest that there are problems with the fit. Poor fits are most likely to be caused by sparse data or by a choice of df that is too low to give an adequate representation of the true trend. If the value of df is less than ~ 4 , the diagnostic plot should be checked particularly carefully for patterns in the residuals, and the df

should be raised if necessary. Sparse data can be detected using the rug plot along the bottom of Fig. 2, which gives the density of observations at each point in time (Chambers and Hastie 1993). The density of observations is high where the rug plot is solid black.

The diagnostic plot in Fig. 2 shows a healthy fit, although standard errors are somewhat higher at the beginning and end of the time period. This is partly due to data sparsity at these points, and also to the fact that the reliability of a GAM fit is always reduced at the end points of the range (Hastie and Tibshirani 1990). If a diagnostic plot indicates that the data are too sparse for the fit to be satisfactory, an alternative model formulation based on covariates might be considered. In that case, the site effects in Eq. 5 would be replaced by predictors derived from habitat variables. Because habitat information can be pooled across all sites, this would remove the need for every site to have its own parameter, and with fewer parameters, the model might provide a better fit to sparse data.

Inference from the GAM indices

Bootstrapped confidence intervals.—Knowledge of the precision with which the GAM trend curve $I(t)$ has been estimated is central to the interpretation of the results. By calculating approximate confidence intervals for I , it may be determined whether or not apparent population changes are statistically significant. In theory, these confidence intervals could be estimated from the standard error estimates provided by S-PLUS for the fitted year effect curve $\hat{s}_d(t)$. However, we prefer to use bootstrap techniques and work directly with the index curve $I(t)$. This avoids the need to transform the standard error estimates for $\hat{s}_d(t)$ into those for $I(t)$, and releases us from assumptions about the statistical distribution of the indices, which would be necessary to convert the standard errors of $I(t)$ into confidence intervals. In addition, it is difficult to obtain accurate estimates of the standard errors of smooth terms in

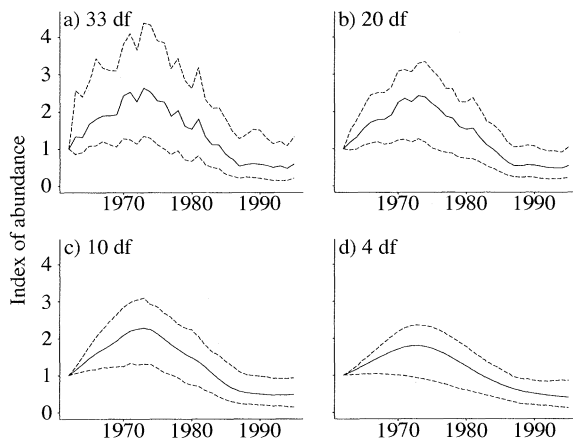


FIG. 3. Bootstrapped 95% confidence limits (dashed lines) for the abundance indices from GAMs with various degrees of freedom applied to CBC data for the Corn Bunting.

GAMs (Chambers and Hastie 1993). The estimates obtained from S-PLUS are adequate for detecting problems with the fit as described in the previous section, but attempts to use them in computing intervals with a precise confidence level are likely to be misleading.

The bootstrap method proceeds by drawing for each of a given number of replicates a random sample of size N , with replacement, from the original N sites. The N sites in the sample are treated as distinct, although in practice there are likely to be some duplicates. A GAM is fitted to the sample from each replicate, and the annual abundance indices are calculated. After B bootstrap replicates, there are B values for the abundance index in any given year. These are sorted into ascending order, and approximate $100(1 - 2\alpha)\%$ confidence limits for that year are provided by the values with rank l and u , where $l = (B + 1)\alpha$ and $u = (B + 1)(1 - \alpha)$; see Buckland (1984).

All results in this paper are obtained from $B = 399$ bootstrap replicates, and this number is likely to be adequate for most applications. Bootstrapping can be a lengthy process, so the number of replicates it is practical to take might be limited by computer resources. When this is the case, compiled software such as the FORTRAN program GAMFIT (Hastie and Tibshirani 1990) might prove useful, as it can provide substantial improvements in speed over the interpreted code of most statistical packages.

Bootstrapped 95% confidence intervals for $I(t)$ are illustrated in Fig. 3 for four different GAMs applied to CBC data for the Corn Bunting. The most notable feature is the narrowing of the confidence interval as the amount of smoothing is increased. This occurs because the estimates of the year effects in models with low levels of smoothing are not constrained by the estimates in adjacent years. Consequently, they are obtained using less information than the estimates from

models with high levels of smoothing, and therefore the uncertainty associated with them is greater.

The narrowing of the confidence intervals demonstrates an important advantage of GAMs over secondarily smoothed indices from other models, in that the abundance trajectory is estimated with better precision. The secondary smoothing of log-linear Poisson regression indices, for example, which correspond to those for the GAM with 33 df, would produce confidence intervals as wide as those in Fig. 3a, because the secondarily smoothed indices would inherit all of the uncertainty of annual fluctuation. Equivalent smoothing inside the model, however, produces narrower confidence intervals as in Fig. 3b, c, d. The improvement can be consequential, as Siriwardena et al. (1998a) found that the confidence intervals associated with secondarily smoothed Mountford abundance indices were often too wide to allow useful inference.

The bootstrap replicates may also be used to find approximate confidence intervals for the percentage population change between any two years. The percentage change is calculated using each of the B index curves; the results are ordered, and the confidence interval is constructed from the lower and upper percentiles as before. Confidence intervals calculated in this way provide approximate hypothesis tests for significant population change. The hypothesis that the percentage population change derives from a distribution with zero mean is rejected at the 5% level if the 95% confidence interval for the change does not contain the point zero. The hypothesis test is two-tailed, and if the distribution of the percentage change satisfies certain symmetry conditions, the approximate test is exact (Buckland 1984).

Analysis of second derivatives.—We now describe a procedure for identifying landmark points in the population trajectory $I(t)$, based on the second derivatives of the curve. The technique is made possible by the smooth differentiable nature of the GAM output, and it provides a useful way of interpreting and comparing the trend curves for different species.

The second derivative of the trend curve $I(t)$ at time t is a measure of the curvature of I at that time. If the second derivative is greater than zero, the curve is turning upward (\cup); if it is less than zero, the curve is turning downward (\cap). The magnitude of the second derivative signifies the tightness of the curvature. Values of approximately zero indicate that the index curve is roughly linear, and the population trajectory is changing at a steady rate. Years in which the second derivative is markedly different from zero are those in which something is happening to alter that rate of population change. Here, we identify years in which the curvature of the index curve is statistically significant, and we refer to these years as *change points*. In conjunction with other temporally referenced data, the timing of the change points might help to suggest causes of the change.

Because I is a nonparametric curve, the second derivatives of I are not available directly as mathematical expressions. Instead, numerical derivative estimates must be used. For example, the first derivative, or gradient, of I at time t can be approximated by

$$\hat{I}'(t) = \frac{I\left(t + \frac{r}{2}\right) - I\left(t - \frac{r}{2}\right)}{r} \quad (8)$$

for some small r . Substituting Eq. 8 into itself, we obtain a simple estimate of the second derivative at point t :

$$\begin{aligned} \hat{I}''(t) &= \frac{\hat{I}'\left(t + \frac{r}{2}\right) - \hat{I}'\left(t - \frac{r}{2}\right)}{r} \\ &= \frac{I(t+r) - 2I(t) + I(t-r)}{r^2}. \end{aligned} \quad (9)$$

In Eq. 9, we use the second differences of I to estimate the second derivatives. More elaborate estimates may be obtained by taking higher differences of I (e.g., Dahlquist and Björck 1974: example 7.5.3). Using the fourth differences, the estimates become

$$\begin{aligned} \hat{I}''(t) &= \frac{1}{12r^2} [-I(t+2r) + 16I(t+r) - 30I(t) \\ &\quad + 16I(t-r) - I(t-2r)] \end{aligned} \quad (10)$$

and to the sixth difference they are

$$\begin{aligned} I''(t) &= \frac{1}{180r^2} \{2I(t+3r) - 27I(t+2r) + 270I(t+r) \\ &\quad - 490I(t) + 270I(t-r) - 27I(t-2r) \\ &\quad + 2I(t-3r)\}. \end{aligned} \quad (11)$$

By placing $r = 1$ in Eqs. 9–11, estimates of the second derivatives are obtained at maximum resolution. In this paper, we use the estimates from the sixth difference, given by Eq. 11, although these must be replaced by the fourth and, finally, the second difference estimates for those years t at which $t \pm 3r$ and $t \pm 2r$ lie beyond the ends of the time series.

Change points, or years in which the second derivative was significantly different from zero, are identified by calculating approximate confidence intervals for the second derivative curve, using the B bootstrap replicates obtained for the abundance indices. The curve of second derivative estimates is calculated for each of the B sets of indices, and the confidence limits are determined by ordering the second derivative point estimates for each year and selecting the appropriate lower and upper percentiles as described previously. The hypothesis that the second derivative estimate for any year derives from a distribution with mean zero is rejected if the confidence interval for the second derivative estimate in that year does not contain the point

zero. From a 95% confidence interval, the two-tailed hypothesis may be rejected at the 5% level.

In years when the second derivative was significantly greater than zero, the population might have undergone either an increase in the rate of growth, or a fall in the rate of decline. Similarly, a second derivative that is significantly negative suggests either a fall in the rate of growth, or an increase in the rate of decline. A significantly positive second derivative therefore indicates a good change for the species, whereas a significantly negative result signifies an adverse change. Occasionally, curvature in the index curve will be nonsignificant, despite appearing to be substantial in magnitude. It is likely that this is caused by uncertainty in the precise timing of the curvature, rather than uncertainty about whether or not substantial curvature has actually occurred.

Covariate models

Extra terms, or *covariates*, may be added to the trend model by including them in the additive predictor of the GAM (Eq. 5). Quantities that can be described by covariates include habitat information such as percentage crop cover or site area, geographical information such as latitude or region, and climatic variables such as average daily minimum temperatures for each year. When sufficient habitat data are available, it might be possible to dispense entirely with the site effect parameters α_i , and replace them with habitat covariates. This has a number of advantages, because separate trends can be formed in every site while the total parametrization is reduced. A low parametrization is particularly important when site abundance records are sparse. In addition, covariate models help to identify which habitat types are beneficial and detrimental to the species.

Each variable in the additive predictor may be either categorical, continuous and parametric, or continuous and nonparametric. For example, in Eq. 5 the site number i is a categorical quantity, whereas t is a continuous variable treated nonparametrically via the expression $s(t)$. When many covariates are included, criteria such as AIC may be used for stepwise model selection (Chambers and Hastie 1993). It is always important to record as many covariates of potential relevance as possible, because serious bias could result if a major source of variation is omitted.

We illustrate covariate models here with a categorical variable representing the geographical region in which each site is located: northeast, northwest, southeast, or southwest. These four regions divide the CBC plots into approximately equal samples around easting 45 and northing 21 of the U.K. national grid (Marchant and Gregory 1994). The regions differ in dominant habitat cover, with the northeast region containing the low-lying, intensively arable land of East Anglia and Lincolnshire, whereas the other regions comprise mainly grazing and mixed farms. Topography, temperatures,

and rainfall patterns also vary between regions. Although the regional model is a simple example of a covariate model, it illustrates how extra terms may be included in the GAM and how their significance may be tested.

We include the region variable through an interaction with the smooth term $s(t)$, with the effect that a separate smooth trend is fitted in each of the four regions. The additive predictor becomes

$$\log(\mu_{it}) = \alpha_i + s_d^{(k)}(t) \quad (12)$$

where μ_{it} is the mean count in site i at time t , and k denotes the region in which site i is located ($k = 1, 2, 3$, or 4). The four smooth regional trends are denoted by $s_d^{(1)}(t)$, $s_d^{(2)}(t)$, $s_d^{(3)}(t)$, and $s_d^{(4)}(t)$, and each have d df. The total parameterization on the smooth terms is therefore increased from d to $4d$. The index curve $I(t)$ is once again the ratio of total predicted abundance (across all regions) in year t to that in year 1.

At present, the fitting of an interaction between a categorical variable and a smooth term is not fully supported in S-PLUS. It is accomplished by inserting four artificial time variables, each with missing values everywhere except on one level of the region variable. An easier alternative for the simple model here is to fit a separate GAM for each region. This would not be possible, however, if the model were to include covariates that were common to all regions.

The significance of the region-by-time interaction may be tested using analysis of deviance techniques (Hastie and Tibshirani 1990). Let D_R be the residual deviance from the regional model, and D_O be the residual deviance from the original model. The statistical distribution of $D_O - D_R$ is approximated by a χ^2 distribution, with degrees of freedom equal to the difference in degrees of freedom between the two models, i.e., $4d - d = 3d$. The region-by-time interaction is considered significant if the observed decrease in deviance due to its inclusion is extreme for the χ^2_{3d} distribution. Although it is known that the difference in residual deviance between two nested GAMs does not have a χ^2 distribution, even asymptotically, the approximation has been shown to provide a useful means of model selection (Hastie and Tibshirani 1990).

The usual Poisson GAM involves a dispersion parameter that is assumed to be unity (Chambers and Hastie 1993). If there is unusually high variance in the data, however, it is common instead to estimate the dispersion parameter within the model. In that case, the test for significant region interaction must be replaced by an F test. Writing ν for the residual degrees of freedom in the regional model (i.e., the number of observations minus the number of parameters), the test statistic is given by $[(D_O - D_R)/3d]/(D_R/\nu)$ and is tested against the $F_{3d,\nu}$ distribution.

RESULTS

The analyses just described are now illustrated using CBC data for 13 British farmland species. The gen-

eralized additive model of Eq. 5 was fitted to the data for each species, using a smoothing spline with 10 df on the year variable. Data records for all species existed for each year between 1962 and 1995. The total number of sites containing records ranged from 185 for the Corn Bunting to 519 for the Chaffinch, whereas the mean number of records per year varied from 35.0 for the Corn Bunting to 93.1 for the Skylark.

For each species, the index curve $I(t)$ was calculated and bootstrapped 95% confidence intervals were determined. Fig. 4 shows the index curves for 12 of the species; results for the Skylark are shown separately in Fig. 5. Diagnostic plots for the 13 GAMs are omitted for brevity, but all were indicative of good fits. At every bootstrap replicate, the second derivative of the index curve was estimated using $r = 1$ in Eq. 11. Significant change points were then identified at the 5% level, and these are also marked on Figs. 4 and 5.

Table 1 shows estimates of percentage population change between selected years for all species. The percentage change between year t_1 and year t_2 is calculated as $(I(t_2) - I(t_1))/I(t_1) \times 100\%$, and approximate 95% confidence limits obtained through bootstrapping are also provided. The change is deemed significant at the 5% level if the confidence interval does not encompass the point 0. Three years were selected for comparison: 1965, 1975, and 1992. Because the GAM fit is less reliable at the end points of the time period, we avoided using the results from 1962 to 1964 at one end and from 1993 to 1995 at the other to ensure that end effects did not bias our inference. The year 1975 was selected as a good intermediate point of comparison because it coincided with a peak in abundance for many of the 13 species.

The results in Table 1 suggest that populations of Bullfinch, Corn Bunting, Grey Partridge, Linnet, Skylark, Tree Sparrow, and Turtle Dove all experienced significant declines on CBC plots between 1965 and 1992. Little overall change was recorded for the Goldfinch, Greenfinch, Reed Bunting, and Yellowhammer, and only the Chaffinch and the Stock Dove saw significant increase between 1965 and 1992. The Yellowhammer is unusual in that its decline began in the mid-1980s rather than in the 1970s. The Stock Dove population underwent a dramatic increase, although exceptionally high variance in the fit reduces the associated significance level. The high variance arises because the observed increase was largely due to substantial changes in only a small number of sites.

The interpretation of the percentage changes in Table 1 must be limited strictly to the sites over which the survey was conducted. To draw conclusions about changes in national abundance, we would have to make the additional assumption that the surveyed plots were representative of the country as a whole. Although land use and cropping patterns on CBC plots have been shown to be representative of those in the lowlands of southern and eastern Britain, they are not typical of the

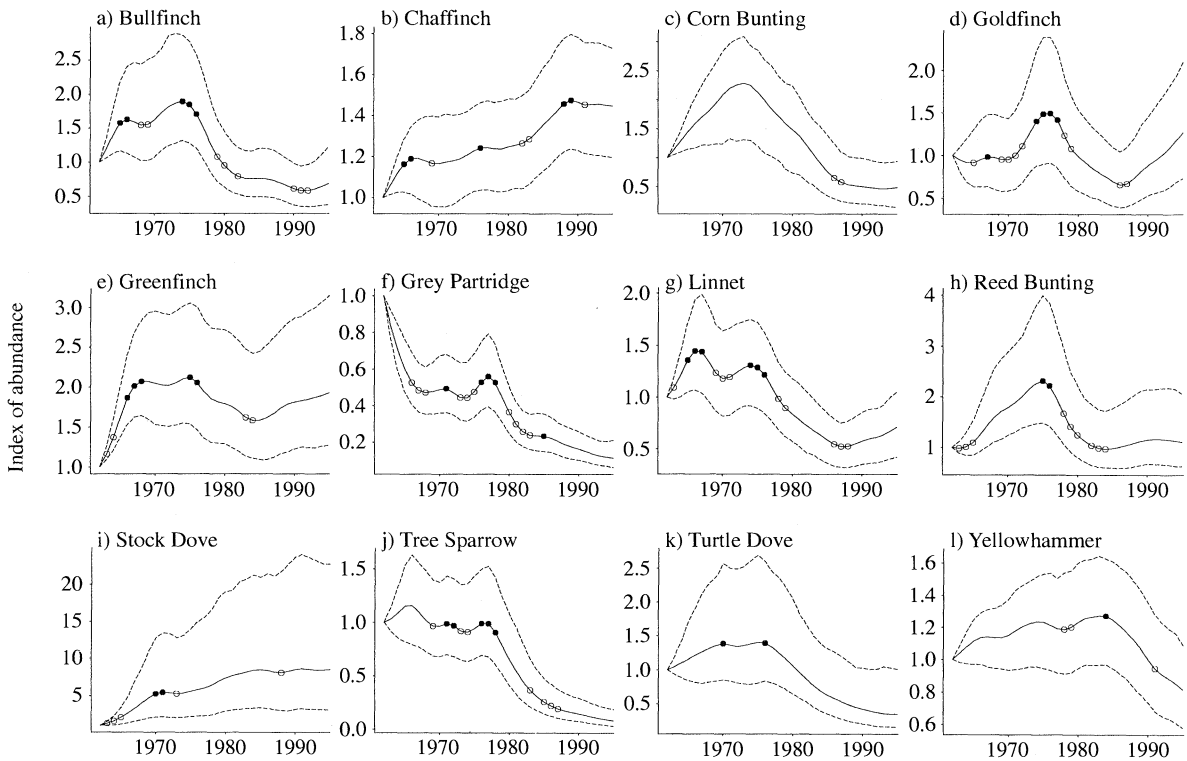


FIG. 4. Index curves from CBC data for 12 farmland species. The solid lines give the index curves from a GAM with 10 df, and the dashed lines represent upper and lower 95% bootstrapped confidence limits. Significant change points are marked on the index curves: a solid circle denotes a point at which the second derivative is significantly negative (a downturn in the index curve), and an open circle denotes a point at which the second derivative is significantly positive (an upturn in the index curve).

landscape in the rest of the country (Fuller et al. 1985). Analysis of CBC data still remains valuable, however, particularly because the southern and eastern regions contain the greater part of the populations of many of the 13 species that we have examined (Gibbons et al. 1993).

Other factors can also complicate the interpretation of estimates of percentage change. One potential problem is the relationship between the number of territory-holding males that were detected in a site and the true number of individuals present in that site. For example, the abundance indices will not reflect the true percentage changes if the probability that a territory is detected varies according to a curvilinear function of local population density (Droegge 1990). Another factor affecting the percentage change estimates is the choice of df used in obtaining the GAM fit. In selecting a value for df, we essentially make an assumption about the form of the trend curve. However, the range of choice for df means that the assumptions behind different GAMs are not as clear as those behind less flexible models. Because these assumptions lead to different criteria in the partitioning of “noise” from “trend,” and because the final abundance estimates are based on trend with noise removed, it is essential to bear in

mind the chosen model when interpreting the estimates of percentage change.

The significant change points marked on the index curves in Fig. 4 reveal some striking patterns, which are more easily visualized in Table 2. From Table 2, it is evident that all species except the Corn Bunting, Stock Dove, and Yellowhammer experienced a significant downturn in population trajectory between 1972 and 1978. In all cases except the Chaffinch, the downturns marked the beginning of a period of decline. Between 1980 and 1989, however, significant upturns in population trajectory occurred for most species. In some cases, these corresponded to the beginning of a period of increase (Chaffinch, Goldfinch, Greenfinch, Linnet, and Reed Bunting), and in others, to a decrease in the rate of decline (Bullfinch, Corn Bunting, Grey Partridge, Skylark, and Tree Sparrow).

The regional covariate model is illustrated using the Skylark data. Independent smooth trend functions were fitted for the four separate regions, each one with 10 degrees of freedom. The results are shown in Fig. 5, with bootstrapped 95% confidence intervals and significant change points also included for each region. The model exposes some noticeable differences between the four regional trends, shown in Fig. 5a–d. In

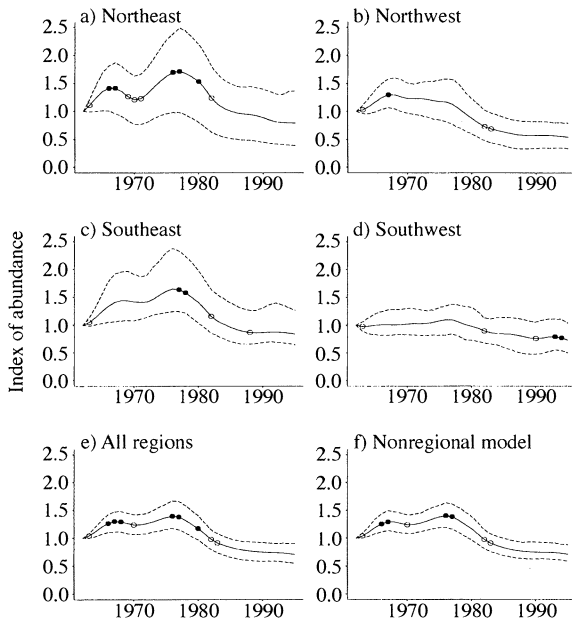


FIG. 5. Results from the region-by-time model for CBC Skylark data. Index curves are shown as solid lines, bootstrapped 95% confidence limits as dashed lines, and change points as solid and open circles. Regional index curves are shown in (a)–(d); in (e), the regional results are pooled to yield overall abundance estimates for the regional model, and in (f), the results of the original (nonregional) model are shown.

particular, the Skylark population appears to have remained almost stable in the southwest, whereas it has undergone clear oscillation in the eastern regions. In all regions, however, a period of stability or increase lasting until the late 1970s has been followed by a period of decline.

The overall abundance indices from the regional model are shown in Fig. 5e. The index for year t is given by the ratio of total predicted count over all re-

gions in year t to that in year 1. One striking feature is the width of the confidence intervals in comparison with those from the regional curves. The narrower confidence intervals arise because the overall indices are based on approximately four times as much information as those from any of the individual regions. The indices from the original model of Eq. 5, without the region-by-time interaction, are shown in Fig. 5f. Because these are almost identical to those from the regional model in Fig. 5e, we deduce that the regional term makes negligible difference to overall inference for the Skylark.

Analysis of deviance tests suggest, however, that the regional model does provide a significant improvement in fit for the Skylark. The observed drop in residual deviance was 180.4 for an increase of 30 equivalent parameters, which is highly significant for the χ^2_{30} distribution. (For significance at the 5% level, a decrease in residual deviance of 43.8 or more is required.) The decision as to whether to use the regional model or the original model therefore hinges on the purpose of the analysis. The regional model highlights discrepant regional trends and gives more accurate predictions at the level of individual sites, but has little bearing on inference about overall trend.

The regional model was also fitted to CBC data for the Chaffinch and Bullfinch. For these species, the regional effect was found to be significant at the 5% level for the Chaffinch (drop in deviance = 178.2), but not for the Bullfinch (drop in deviance = 39.9). In the case of the Chaffinch, although an overall increase occurred in all regions, the increase was steady in the west, whereas the eastern regions saw a fall in abundance between 1972 and 1980. The pattern of change for the Bullfinch was broadly similar across all regions.

DISCUSSION

Comparison of the GAM approach with other methods

The GAMs described in this paper belong to the class of sites-by-years models for survey data, of which log-

TABLE 1. Estimated percentage population changes for 13 farmland bird species using output from GAMs with 10 df fitted to CBC data, together with bootstrapped 95% confidence intervals (in parentheses).

Species	Percentage population change (95% CI)		
	1965–1975	1975–1992	1965–1992
Bullfinch	17 (–1, 38)	–68* (–78, –58)	–63* (–76, –49)
Chaffinch	6 (–4, 16)	19* (11, 28)	25* (11, 41)
Corn Bunting	47 (–13, 82)	–78* (–90, –57)	–67* (–87, –31)
Goldfinch	63* (25, 112)	–33* (–47, –14)	9 (–23, 57)
Greenfinch	31* (5, 63)	–13 (–34, 15)	14 (–19, 58)
Grey Partridge	–21* (–35, –1)	–71* (–79, –59)	–77* (–85, –66)
Linnet	–5 (–18, 9)	–52* (–65, –37)	–55* (–68, –40)
Reed Bunting	111* (56, 192)	–50* (–66, –25)	6 (–34, 73)
Skylark	16* (2, 31)	–46* (–52, –38)	–37* (–46, –25)
Stock Dove	169* (57, 363)	53* (16, 114)	311* (115, 704)
Tree Sparrow	–18 (–37, 13)	–87* (–94, –78)	–90* (–95, –80)
Turtle Dove	21 (–18, 84)	–73* (–86, –37)	–67* (–83, –46)
Yellowhammer	11 (–9, 29)	–26* (–37, –14)	–18 (–36, 4)

Notes: The percentage change is obtained as $100(I_2 - I_1)/I_1$, where I_1 is the index for the earlier year and I_2 is that for the later year. Changes that are significant at the 5% level are marked with an asterisk.

TABLE 2. Significant change points for the 13 farmland species from 1963 to 1994. Solid circles denote significant downturns in the population trajectory; open circles denote significant upturns.

Species	1963	1965	1967	1969	1971	1973	1975	1977	1979	1981	1983	1985	1987	1989	1991	1993	
Bullfinch		●	●	○	○		●	●	●	○	○				○	○	○
Chaffinch		●	●	○				●				○	○				
Corn Bunting																	
Goldfinch		○	●	○	○	○	○	●	●	●	○	○					
Greenfinch	○	○	●	●	●			○	○								
Grey Partridge			○	○				○	○	○	○	○	○	○	○	○	○
Linnet	○	●	●	●	○	○	○	○	○	○	○	○	○	○	○	○	○
Reed Bunting	○	○	○				●	●	○	○	○	○	○	○	○	○	○
Skylark	○	○	●	●	○			●	●			○	○				
Stock Dove	○	○			●	●	○							○			
Tree Sparrow				○	●	●	○	○	●	●	●		○	○	○	○	○
Turtle Dove				●				●									
Yellowhammer									○	○		●					○

linear Poisson regression (ter Braak et al. 1994, Pannekoek and van Strien 1996) and the Mountford method (Mountford 1982, 1985, Peach and Baillie 1994, ter Braak et al. 1994) are also members. Of these alternatives, the GAM approach is that best suited to the modeling of long-term nonlinear trends; once the complexity of trend has been specified, the GAM will provide the optimal fit to the data. Secondary smoothing of the output obtained from log-linear Poisson regression and the Mountford method does not provide optimal fit, and produces abundance estimates with lower precision.

The software package TRIM (Pannekoek and van Strien 1996), which is especially designed to fit log-linear Poisson regression models to wildlife census data, also incorporates an additional feature to model the within-site correlation structure between counts in different years. This autocorrelation correction is designed to improve estimates of the standard errors of the model indices, while usually having a negligible effect on the estimates of the model indices themselves (Pannekoek and van Strien 1996). An alternative solution to the problem of variance estimation when there is serial correlation in the counts is provided by the bootstrap methods used in this paper, which also remove the need to assume a parametric model for the correlation structure. The drawback is that bootstrapping can be very computer intensive.

The route regression approaches used for the North American Breeding Birds Survey (Droege 1990) have recently been enhanced to allow both nonlinear route-specific trends and inclusion of covariates (James et al. 1990, 1996). Scatterplot smoothers are used to provide smooth temporal abundance trends in each route. The models are not formulated as GAMs because the structure of error distribution and link function is missing. Recasting them into a GAM framework, however, would be very straightforward.

The principal difference between the enhanced route regression method and the GAM approach lies in the

calculation of overall trend at the national or regional scale. In route regression, the disparate route-specific trends are aggregated into an overall trend using various weighting procedures. The weight allotted to a particular route is subjectively chosen, often based on area covered or precision of trend estimate along the route. The choice of weighting scheme can have a marked effect on the final trend estimate (James et al. 1990). By contrast, the overall index from the GAM approach at any time is proportional to the total predicted abundance at that time, over all sites surveyed. This index has intuitive appeal, especially if the surveyed sites are representative of the whole country or region.

Other recent developments in route regression include the models of Link and Sauer (1997a, b). In these, temporal trend is represented as a polynomial function of time, and the models are formulated as GLMs with logarithmic link function and Poisson or negative binomial distributions for the counts. The smooth trend curve $s(t)$ of Eq. 5 is therefore constrained to a polynomial shape. In concept, the models are similar to the GAM formulation of this paper, although the parametric trend curve is badly behaved when extrapolated beyond the limits of the time series. Application of the polynomial model with Poisson error distribution to the CBC data suggested that the indices from a polynomial fit of order 10 were very close to those obtained from a GAM with 10 df. The two models have equivalent parameterization, and both gave similar results for residual deviance. In our view, it is preferable to avoid placing parametric constraints on the shape of the trend curve, although the similarity of the results obtained from the two methods is encouraging.

The basic GAM detailed in Eq. 5 is simplistic in assuming that expected abundance follows the same trend across all sites. Within the GAM framework, however, there is much potential to go beyond the basic model. With site effects replaced by covariates such as habitat, geographical, and climate variables, it would

be possible to obtain a separate trend in each site while using far fewer parameters in the model. If required, weighting of the contributions from different sites could also be incorporated. Given sufficient environmental data, therefore, the GAM approach can embody all the advantages of route regression, and can also be fitted to data that are too sparse for route regression techniques. In the CBC, for example, many sites were not surveyed for a sufficiently long time period to enable site-specific regressions to be performed.

The final decision as to the most appropriate analysis method rests on the particular data available and the geographical scale involved. It is worth noting that the methods presented in this paper for inference from the GAM index curve may be applied equally well to the output from other models. These enable estimation of the significance of changes in abundance to be performed without using tests that are based on linear trends (e.g., James et al. 1996), and power analyses (Gerrodette 1987, 1991).

Extensions

The GAM formulation presented in this paper involves a Poisson error distribution and a logarithmic link function. Other choices for both error distribution and link function are available (Hastie and Tibshirani 1990). In particular, there has been recent interest in using the negative binomial distribution in place of the Poisson for environmental data, as it is thought to cope better with overdispersion (Gotway and Stroup 1997, Link and Sauer 1997a). The logarithmic link function is a popular choice to use with this error distribution.

We have indicated in several places the advantages of adding covariates to the basic GAM: these may enter the additive predictor as linear, nonlinear, or categorical variables, and allow environmental effects such as rainfall or habitat to be included in abundance modeling. This might be one of the most promising lines of future enquiry. In order for full covariate models to be successful, all variables of potential importance must be identified at the outset of a monitoring scheme and recorded throughout the survey period.

Monitoring schemes should be designed to cover a representative sample of sites: otherwise, the trends obtained must be interpreted strictly as trends over selected habitat rather than true national trends. Continuity of survey cover is also important; continuous coverage of a small sample of sites is more valuable for trend analysis than patchy coverage of a larger sample. These issues have been addressed in the U.K. Breeding Bird Survey (Gregory et al., *in press*), which is in the process of replacing the CBC.

When the survey covers a random sample of sites, generalized additive mixed models (GAMMs) might be worth considering. Site effects in a GAMM are no longer regarded as fixed unknowns, as in Eq. 5, but as realizations from a parametric distribution whose parameters are to be estimated. Without a random sample

of sites, the data could lie predominantly in the tail of this distribution, causing bias in prediction. The mixed model would be more parsimonious than the fixed-effects model, although some investigation of a suitable trend index would be required. Although there has been much recent interest in methodology for generalized linear mixed models (e.g., Schall 1991, Engel and Keen 1994), we are not aware of any work to date on generalized additive mixed models. Development of such models might prove a valuable tool for estimating trends in wildlife populations from annual surveys of randomly selected sites.

Results for British seed eaters

The GAM index curves in Fig. 4 show various periods of decline, increase, and stability for the farmland species. Together with further data giving demographic rates over the contrasting time periods, the results help to suggest mechanisms underlying the population change (e.g., Siriwardena et al. 1998b). Additional clues about the causes of change are provided by different species whose trend curves share common characteristics. Examination of ecological similarities and differences between the species highlight environmental factors that could be responsible.

The regional model (Eq. 12) applied to the Skylark data reveals interesting differences in regional trend (Fig. 5). The steeper declines in the east might reflect the arable farmland that is characteristic of these regions. Arable areas have been more dramatically affected than grazing lands by some aspects of agricultural change over the last 30 years, and also the baseline population density in the early 1960s was lower in arable regions. In the mid-1960s, populations increased in all regions except the southwest. These increases probably represent population recoveries following the harsh winter weather of 1962–1963, which was less severe in the southwest region. Use of the regional model is encouraged for all species, at least for exploratory data analysis.

The change points shown in Table 2 reveal conspicuous patterns across species. The mixed fortunes of species in the mid-1960s probably reflect differing responses to the harsh winter weather of 1962–1963. Depending on ecology and demography, by 1965–1966, most affected species were either reaching the end of a period of steep decline or arriving at a plateau after a steep recovery (Baillie 1990, Marchant et al. 1990, Greenwood and Baillie 1991). The 1970s saw a range of changes in agriculture that are thought to have been detrimental to farmland bird populations (O'Connor and Shrubbs 1986, Fuller et al. 1995, Baillie et al. 1997, Siriwardena et al. 1998a). This is borne out by the collection of negative change points observed at that time.

Some variation between species in the precise timing of change points is to be expected, due to differences in life-span and reproductive strategy. For example,

coincident depressions in breeding success for a short-lived and a long-lived species tend to be manifested in the abundance of the short-lived species much earlier, whereas concurrent depressions in adult survival are apparent first in the abundance of the long-lived species. This is especially true if individuals of the short-lived species mature earlier and produce more offspring than those of the long-lived species.

Although downturns dominated the 1970s, the reverse is true of the 1980s. For the most part, the upturns of the 1980s have been followed by a period of relative stability, albeit at a lower level than prior to the mid-1970s. The farmland populations of the 13 seed-eating species have therefore failed to recover from their severe declines.

The significant change points identified from the GAM index curves in Fig. 4 may be compared against those obtained from secondarily smoothed Mounford indices by Siriwardena et al. (1998a). The change points found by the two methods correspond fairly closely, as do the smoothed abundance curves. The GAM approach is to be favored because of its better precision, but the close correspondence with the results from the Mounford method indicates that earlier analyses need not be discarded.

Prospects

In conclusion, we believe that the GAM approach provides the most general and flexible framework currently available for the analysis of trends in census data. With the CBC data, there are encouraging similarities between GAM indices and those obtained from log-linear Poisson regression and the Mounford method, although inference based on more primitive techniques such as the chain method should be reassessed. The GAM framework may also be adapted to provide improved analyses for other national bird censuses and wildlife monitoring schemes, which are currently indexed using alternative methods (ter Braak et al. 1994, Thomas 1996).

Recent improvements in the design of monitoring surveys will enable more comprehensive analyses to be carried out in the future. The new U.K. Breeding Bird Survey, for example, has been designed to cover a stratified random sample of survey grid squares (Gregory et al., *in press*). A further U.K. bird monitoring scheme uses mark-recapture data to index abundance, and is currently analyzed using generalized linear models (Peach et al. 1998). This may easily be extended into a GAM framework. There is also considerable potential for the application of GAMs in models for survival and breeding success, and work is currently underway to analyze daily nest failure rates in this manner.

ACKNOWLEDGMENTS

We would like to thank Raymond O'Connor and an anonymous referee for helpful comments on the first draft. This work was funded partly by the U.K. Ministry for Agriculture,

Fisheries and Food, under contract CSA3109, and partly through a research studentship from the U.K. Engineering and Physical Sciences Research Council (EPSRC). The CBC has been funded by the U.K. Department of the Environment, and by a contract with the Joint Nature Conservation Committee on behalf of the Countryside Council for Wales, English Nature, Scottish National Heritage, and the Department of the Environment for Northern Ireland. We are indebted to the volunteers who have collected CBC data since the inception of the census, and to John Marchant, Andrew Wilson, and previous CBC analysts for interpreting and collating the field data.

LITERATURE CITED

- Baillie, S. R. 1990. Integrated population monitoring of breeding birds in Britain and Ireland. *Ibis* **132**:151–166.
- Baillie, S. R., R. D. Gregory, and G. M. Siriwardena. 1997. Farmland bird declines: patterns, processes and prospects. *Biodiversity and Conservation in Agriculture. British Crop Protection Council Symposium Proceedings* **69**:65–87.
- Barker, R. J., and J. R. Sauer. 1992. Modelling population change from time series data. Pages 182–194 in D. R. McCullough and R. H. Barrett, editors. *Wildlife 2001: populations*. Elsevier, London, UK.
- Bowerman, B. L., and R. T. O'Connell. 1987. *Time series forecasting*. Second edition. Duxbury Press, Boston, Massachusetts, USA.
- Buckland, S. T. 1984. Monte Carlo confidence intervals. *Biometrics* **40**:811–817.
- Chambers, J. M., and T. J. Hastie. 1993. *Statistical models in S*. Chapman and Hall, London, UK.
- Dahlquist, G., and Å. Björck. 1974. *Numerical methods*. Prentice Hall, Englewood Cliffs, New Jersey, USA.
- Droege, S. 1990. The North American Breeding Bird Survey. Pages 1–4 in J. R. Sauer and S. Droege, editors. *Survey designs and statistical methods for the estimation of avian population trends*. U.S. Fish and Wildlife Service Biological Report **90**(1).
- Engel, B., and A. Keen. 1994. A simple approach for the analysis of generalized linear mixed models. *Statistica Neerlandica* **48**:1–22.
- Francis, B., M. Green, and C. Payne, editors. 1993. *The GLIM system: release 4 manual*. Oxford University Press, Oxford, UK.
- Fuller, R. J., R. D. Gregory, D. W. Gibbons, J. H. Marchant, J. D. Wilson, S. R. Baillie, and N. Carter. 1995. Population declines and range contractions among lowland farmland birds in Britain. *Conservation Biology* **9**:1425–1441.
- Fuller, R. J., J. H. Marchant, and R. A. Morgan. 1985. How representative of agricultural practice in Britain are Common Birds Census farmland plots? *Bird Study* **32**:56–70.
- Geissler, P. H., and B. R. Noon. 1981. Estimates of avian population trends from the North American Breeding Bird Survey. Pages 42–51 in C. J. Ralph and J. M. Scott, editors. *Estimating numbers of terrestrial birds*. *Studies in Avian Biology* **6**.
- Geissler, P. H., and J. R. Sauer. 1990. Topics in route regression analysis. Pages 54–57 in J. R. Sauer and S. Droege, editors. *Survey designs and statistical methods for the estimation of avian population trends*. U.S. Fish and Wildlife Service Biological Report **90**(1).
- Gerrodette, T. 1987. A power analysis for detecting trends. *Ecology* **68**:1364–1372.
- Gerrodette, T. 1991. Models for power of detecting trends—a reply to Link and Hatfield. *Ecology* **72**:1889–1892.
- Gibbons, D. W., J. B. Reid, and R. A. Chapman. 1993. *The new atlas of breeding birds in Britain and Ireland, 1988–1991*. Poyser, London, UK.
- Gotway, C. A., and W. W. Stroup. 1997. A generalized linear model approach to spatial data analysis and prediction.

- Journal of Agricultural, Biological and Environmental Statistics **2**:157–178.
- Greenwood, J. J. D., and S. R. Baillie. 1991. Effects of density dependence and weather on population changes of English passerines using a non-experimental paradigm. *Ibis* **133**(suppl. 1):121–133.
- Gregory, R. D., S. R. Baillie, and R. I. Bashford. *In press*. Monitoring breeding birds in the United Kingdom. *In Proceedings of the 13th International Conference of European Bird Census Councils*. Estonia 1995.
- Hastie, T. J., and R. J. Tibshirani. 1990. Generalized additive models. Chapman and Hall, London, UK.
- James, F. C., C. E. McCulloch, and D. A. Wiedenfeld. 1996. New approaches to the analysis of population trends in land birds. *Ecology* **77**:13–27.
- James, F. C., C. E. McCulloch, and L. E. Wolfe. 1990. Methodological issues in the estimation of trends in bird populations with an example: the Pine Warbler. Pages 84–97 *in* J. R. Sauer and S. Droege, editors. Survey designs and statistical methods for the estimation of avian population trends. U.S. Fish and Wildlife Service Biological Report **90**(1).
- Link, W. A., and J. R. Sauer. 1997a. Estimation of population trajectories from count data. *Biometrics* **53**:488–497.
- Link, W. A., and J. R. Sauer. 1997b. New approaches to the analysis of population trends in land birds: comment. *Ecology* **78**:2632–2634.
- Marchant, J. H., and R. D. Gregory. 1994. Recent population changes among seed-eating passerines in the United Kingdom. Pages 87–95 *in* W. Hagemeijer and T. Verstrael, editors. Bird Numbers 1992. Distribution, Monitoring and Ecological Aspects. Proceedings of the 12th International Conference of the International Bird Census Committee and European Ornithological Atlas Committee. SOVON, Beek-Ubbergen, The Netherlands.
- Marchant, J. H., R. Hudson, S. P. Carter, and P. Whittington. 1990. Population trends in British breeding birds. British Trust for Ornithology, Tring, Hertfordshire, UK.
- McCullagh, P., and J. A. Nelder. 1989. Generalized linear models. Second edition. Chapman and Hall, London, UK.
- Mountford, M. D. 1982. Estimation of population fluctuations with application to the Common Birds Census. *Applied Statistics* **31**:135–143.
- Mountford, M. D. 1985. An index of population change with application to the Common Birds Census. Pages 121–132 *in* B. J. T. Morgan and P. M. North, editors. Statistics in ornithology. Springer-Verlag, Berlin, Germany.
- Nelder, J. A., and R. W. M. Wedderburn. 1972. Generalized linear models. *Journal of the Royal Statistical Society Series A* **135**:370–384.
- O'Connor, R. J., and M. Shrubbs. 1986. Farming and birds. Cambridge University Press, Cambridge, UK.
- Pannekoek, J., and A. van Strien. 1996. TRIM (Trends and Indices for Monitoring data). Research paper 9634. Statistics Netherlands, Voorburg, The Netherlands.
- Peach, W. J., and S. R. Baillie. 1994. Implementation of the Mountford indexing method for the Common Birds Census. Pages 653–662 *in* W. Hagemeijer and T. Verstrael, editors. Bird Numbers 1992. Distribution, Monitoring and Ecological Aspects. Proceedings of the 12th International Conference of the International Bird Census Committee and European Ornithological Atlas Committee. SOVON, Beek-Ubbergen, The Netherlands.
- Peach, W. J., S. R. Baillie, and D. E. Balmer. 1998. Long-term changes in the abundance of small passerines in Britain and Ireland as measured by constant effort mist-netting. *Bird Study* **45**:257–275.
- SAS Institute. 1996. SAS/STAT software: changes and enhancements through release 6.11. SAS Institute, Cary, North Carolina, USA.
- Sauer, J. R., and P. H. Geissler. 1990. Annual indices from route regression analyses. Pages 58–62 *in* J. R. Sauer and S. Droege, editors. Survey designs and statistical methods for the estimation of avian population trends. U.S. Fish and Wildlife Service Biological Report **90**(1).
- Schall, R. 1991. Estimation in generalized linear models with random effects. *Biometrika* **78**:719–727.
- Siriwardena, G. M., S. R. Baillie, S. T. Buckland, R. M. Fewster, J. H. Marchant, and J. D. Wilson. 1998a. Trends in the abundance of farmland birds: a quantitative comparison of smoothed Common Birds Census indices. *Journal of Applied Ecology* **35**:24–43.
- Siriwardena, G. M., S. R. Baillie, and J. D. Wilson. 1998b. Variation in the survival rates of British passerines with respect to their population trends on farmland. *Bird Study* **45**:276–292.
- Statistical Sciences. 1993. S-PLUS reference manual. StatSci, a division of Mathsoft, Seattle, Washington, USA.
- Taub, S. R. 1990. Smoothed scatterplot analysis of long-term Breeding Bird Census data. Pages 80–83 *in* J. R. Sauer and S. Droege, editors. Survey designs and statistical methods for the estimation of avian population trends. U.S. Fish and Wildlife Service Biological Report **90**(1).
- ter Braak, C. J. F., A. J. van Strien, R. Meijer, and T. J. Verstrael. 1994. Analysis of monitoring data with many missing values: which method? Pages 663–673 *in* W. Hagemeijer and T. Verstrael, editors. Bird Numbers 1992. Distribution, Monitoring and Ecological Aspects. Proceedings of the 12th International Conference of the International Bird Census Committee and European Ornithological Atlas Committee. SOVON, Beek-Ubbergen, The Netherlands.
- Thomas, L. 1996. Monitoring long-term population change: why are there so many analysis methods? *Ecology* **77**:49–58.
- Tucker, G. M., and M. F. Heath. 1994. Birds in Europe: their conservation status. BirdLife International, Cambridge, UK.