Towards a method for the estimation and use of averaged multi-species trends, as indicators of patterns of change in butterfly populations.

Stephen Freeman, Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB.

March 2009

Introduction

The UK Butterfly Monitoring Scheme (UKBMS).

Since its establishment in 1976 and subsequent expansion, the UK Butterfly Monitoring Scheme has accumulated records from up to 26 weekly visits each year to a large number of survey sites (almost 1700 sites in total, with almost 800 sites currently active). The weekly counts are summed each year to give an annual total for each species at each site. The resulting data provide the best source of information on annual population changes in most widespread species and, as the sites are selected by the surveyors (rather than via a formally randomised design), even for several species with a localised distribution. Among common species, only elusive, canopy-dwelling species are relatively poorly represented. Analyses are provided in a series of annual reports (most recently Botham et al., 2008) and a number of scientific papers (e.g. Brereton et al. 2008, Roy et al. 2001)

Analytical methods to date

Models for site-specific annual survey data.

The structure of the UKBMS survey, comprising annual counts at a large number of sites, is similar in broad outline to that long-used for other common and widespread taxonomic groups, e.g. breeding birds (Marchant et al., 1990; Fewster et al., 2000; Freeman et al., 2007). Estimation of annual trends in the UKBMS therefore has also followed methods used for other taxa, based on Poisson models (on account of the non-negative integer status of the data) with 'Site' employed as a multi-level additive factor (on a logarithmic scale). This naturally accounts for the varying numbers of butterflies at sites which vary greatly in suitability for a species but permits these to vary in constant proportion between years.

The most straightforward models also permit 'Year' to be a multi-level factor; trends at each site move in parallel but no constraint is otherwise imposed upon annual changes. This, a 'Generalized Linear Model' (GLM: Nelder and Wedderburn, 1972; McCullagh and Nelder, 1989), is now easily fitted in widely-available software and has the additional advantage that site turnover and sporadic years missed are easily accommodated (see e.g. ter Braak et al, 1994). Under such models the expected value of the total count C_{it} for a species at site i in year t as a function of these site (S_i) and year (Y_t) effects is simply:

(1) $$\text{Log}(\ \text{E}(C_{it})\ ) = S_i + Y_t$$

Inverting the link function, imposed to keep fitted values positive, e^{Yt} is then adopted as an index of the species' relative abundance in year t. An arbitrary constraint on the parameters in (1) is required for identifiability; frequently, the first year effect is set to zero, making trends for different species immediately comparable. The GLM modelling framework provides great capacity for statistical inference and is highly flexible; model (1) is readily extended to include a weekly effect W_k and fitted to data expressed at a finer temporal resolution, or constrained to impose relationships between annual indices and external variables, e.g. a simple linear trend (constant rate of change between consecutive years) or a dependence upon environmental conditions such as temperature or rainfall.

Despite this flexibility, however, due to the high year-to-year variability in butterfly numbers, annual indices based on (1) are highly volatile, and enduring trends difficult to identify amongst more ephemeral, possibly weather-driven fluctuations. Imposing a simple linear trend removes this extraneous variability but is an unrealistic model in practice – trends in nature rarely progress at a constant rate for any length of time. Furthermore, as a means of identifying species in decline, a linear model is a blunt tool; consider a species increasing before going into decline – many years of decline are likely to be required before the positive slope initially produced by such a model finally becomes negative.

For this reason, a means of 'smoothing' trends in a less restrictive fashion is now adopted. Fewster et al (2000) first introduced a Generalized Additive Model (GAM: Hastie and Tibshirani, 1990) for bird census data. A GAM replaces the independent annual effects Y_t in (1) with one of a number of available forms of a smooth function $f(Y_t)$. The model is nonparametric and, although harder to fit than a simple GLM (1) – itself a special case of GAM – several widely-used packages now have this capacity. The extent of smoothing produced is set by a user-defined number of 'degrees of freedom (d.f.)' (see Hastie and Tibshirani 1990), with fewer d.f. producing greater smoothing. The amount of smoothing can therefore be set empirically, a compromise between too little (which retains excessive year-to-year variability) and too much (which irons out longer-term changes in trend direction that may be of interest). For avian data, degrees of freedom numbering some 30-40% of the number of years in the survey have been found most useful (Fewster et al., 2000) and adopted in butterfly work to date. An obvious parametric alternative, constraining the year effects in (1) by a polynomial of appropriate order, is not favoured as model behaviour towards the ends of the series, which are especially crucial in determining the current status of a species, can be erratic. An application of GAM to British butterfly data is provided by Rothery and Roy (2001).

Multi-species 'biodiversity' Indicators

Given a set of single-species temporal trends calculated as above, it has become increasingly common to produce a single composite 'indicator' of fluctuations representative of a group of species. That such indicators are of a high profile in policy making is evidenced by the adoption of a farmland bird indicator (Freeman et al., 2001) into the UK government's 'Quality of Life' index, and a commitment to reverse the decline in the species concerned by 2020. A similar indicator scheme for butterflies was established in 2005, and subsequently extended to cover separately butterflies in England and Scotland, and habitats (farmland and woodland) in England.

Indicators are constructed by firstly estimating individual species' trends as above and calculating the geometric means across contributing species of each year's index values. The geometric mean (based on an arithmetic mean of the log-transformed species' indices) means that the combined effects of, say, one species halving in number and another doubling cancel one another out and leave the averaged index unchanged. This feature makes the geometric mean more appealing than a simple arithmetic mean of relative abundances in consecutive years which, in the previous example, equates to an indicator increasing by 25%, with positive connotations that are not justified by the data. Freeman et al. (2001) found little difference whether single-species indices were smoothed prior to averaging, or the average trajectory derived from unsmoothed species trends was smoothed subsequently. In practice, the former method is usually adopted though we shall later consider the latter in this report.

Having produced a multi-species indicator, interest has centred upon whether changes over set periods of time are statistically significant, with a view to determining whether remedial measures have proved successful. Previous construction of indicators has found that reliable confidence intervals (CI) require a bootstrapping (resampling) procedure, selecting an appropriate number of sites at random, and with replacement, from those in the data and refitting the GAM to the data for those sites. This process is computationally straightforward, but time-consuming where large amounts of data are involved. Such confidence limits can be employed to test for changes in the indicator between consecutive (or indeed arbitary) pairs of years (non-overlapping 85% CI indicating a significant difference, at α=0.05, between two indices: Anganuzzi, 1993). The bootstrapped replications can also be used to set CI directly upon the estimated changes between consecutive years (Freeman et al., 2001) or, by incorporating numerical approximations to the second derivatives, to identify those years in which a significant change in direction occurred (Siriwardena et al., 1998). All of these procedures are applicable equally to single- or to multi-species trends.

Alternative approaches to (a) smoothing of trends and (b) quantifying conservation status.

Given the computationally intensive nature of any inferential methods based upon bootstrapping non-parametric statistical models, attention has recently shifted towards the consideration of more tractable alternatives.

In the 2007 annual UKBMS report (Botham et al., 2008; p.18) the use of butterfly indicators was discussed based upon smoothing an aggregate index post-hoc using a state-space model (SSM). SSM are characterised by having two components, a *process* component in which a state of interest at a given time, here an index of the abundance of one or more species, is functionally (but stochastically) related to the state at the previous time interval(s). That is, in expectation the abundance index in year *t* is given by some function *T* of that in the previous year(s), and an appropriate distribution is assumed for the error term. The second part of the model is the *observation* equation, which assumes that our measurement of such an index is itself subject to a degree of error around the true value. Unlike a standard GLM/GAM then, two sources of error are identified, that arising from the stochastic model assumed for the population changes, and the observation error associated with imperfect measurement (via survey sampling error) of the population level. The state-space form permits estimation both of the variances of these error terms (rarely of interest in themselves) and any parameters established in the model *T*.

A SSM therefore is parametric (unlike a GAM) but it serves also, via *T*, to smooth a series, the extent of smoothing determined by the relative estimated variances of the two sources of error. Statistical theory and hence formulae for standard errors and confidence limits are well-established. Unlike a GAM, it is not required to set d.f. for smoothing beforehand, the degree of smoothing is determined only by the data and the model selected. The latter is likely to be empirical, although in well-monitored taxa with high-quality demographic information also available the model can be biologically informed, or directly process-based, e.g. with *T* taking the form of a Leslie matrix familiar in population ecology (Mazzetta et al., 2007; Brooks et al., 2008).

Initial SSM indicators for butterflies (Botham et al., 2008) were derived by producing a multi-species indicator from (unsmoothed) species' indices and then smoothing the resulting series. Normal errors were assumed for both process and observation errors and the model fitted by maximum likelihood using the Kalman Filter (see Appendix) and the software TRENDSPOTTER (Visser, 2004). This package is also being used in the construction of similar indicators for birds (Gregory et al., 2007, 2008). Although in principle the Kalman filter can be used to model vectors of observations (of the kind arising in records from a large number of sites), in short series of data computational problems can arise making this impractical (Besbeas et al., 2005). Initially, further explorations have therefore been based upon single series via this package and one of the models provided, which relates consecutive values of the true index μ_t thus:

(2) $$\mu_t - 2\mu_{t-1} + \mu_{t-2} = \eta_t$$

where η_t is normally-distributed error with zero mean and constant variance. Introducing a second variable $\lambda_t = -\mu_{(t-1)}$ permits, after some algebra, the rewriting in matrix form:

(3) $$\begin{pmatrix} \mu_{t+1} \\ \lambda_{t+1} \end{pmatrix} = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix} \begin{pmatrix} \mu_t \\ \lambda_t \end{pmatrix} + \begin{bmatrix} \eta_{t+1} \\ 0 \end{bmatrix}$$

thus the state vector $(\mu, \lambda)^T$ in any given year is a linear function of that in the previous year, permitting the use of the Kalman filter in likelihood construction, with transition matrix *T* (see Appendix):

$$T = \begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$$

A simple application of this model to an unsmoothed index from UKBMS data for an example set of species {species covered by the wider countryside indicator; Botham et al., 2008} is shown in Figure 1. Little difference was found between the Integrated Random Walk model (introduced above) or a local linear trend model (Harvey, 1989). The IRW model is adopted hereafter. The smoothed trend clearly captures the general trend, but is perhaps oversmoothed for the purposes of retaining key features while identifying significant declines; standard deviations are small. We note (see Appendix, and Visser, 2004) that the likelihood is formed for data omitting a number of values early in the series, to initialise the Kalman filter. TRENDSPOTTER suggests about 10 data points be omitted in this way, but this is a substantial percentage of the total in this case (Visser, 2004, used a much longer time series). This limits the precision with which error variances are estimated for the UKBMS trend and suggests most of the variability is attributed to the observation process. Reducing the number of data points to about five we have found entirely adequate, the smoothed trend is little affected and the variances are increased, though not to any great extent (Figure 1).

Nonetheless, we consider the SSM framework to be one of enormous potential in butterfly monitoring though further models and methods of variance estimation/partitioning could beneficially be explored. To this end, we have developed code, yet to be fully validated, written in the statistical package R and utilising a completely general form for the transition matrix *T*. This will permit the fitting of arbitrary linear state-space models to smooth the data in addition to those offered by TRENDSPOTTER, one or more of which one may be more suitable for the task in question. Custom-built R programmes will also provide us with the means for efficiently integrating the process within the annual updating of analyses of UKBMS data. We note too that a consequence of the methods employed to estimate the initial (unsmoothed) trends, at either a single- or multi-species level, permit the estimation of associated estimates of precision. Such estimates can then also be transferred to the Kalman filter model, in which the observation error variances are currently held constant and estimated entirely freely. The addition of prior estimates of observation error to the Kalman filter algorithm this way is not offered by TRENDSPOTTER, but will be feasible in our own programme, and might improve the estimation of process error variance with such limited data, and is a subject of ongoing research.

SSM based on non-normal error assumptions can also be fitted, most easily in a Bayesian framework (Brooks et al., 2008), and this too may prove a profitable area to investigate further. Analyses of data at the site level (rather than of a single index derived from them) is also easier in this context, where the Kalman filter is not required.

We note finally that conservationists are increasingly turning not to the identification of a trend or change in numbers as significantly different from zero, or not, but to ‘equivalence testing’. Here, the traditional (point) null hypothesis of no change, {H0: β=0}, where β is, say, the slope of a regression line or some other indicator of ‘change’ is replaced by the hypothesis {H0: U > β> L}, i.e. that β lies between limits U and L, a range which is determined *a priori* to represent that within which the change is ecologically insignificant, or not requiring of current conservation concern (see e.g. Camp et al., 2008). SSM and/or Bayesian methods of model-fitting provide a natural context for such an analysis – see Brooks et al (2008).

Recommendations

Based on this assessment we recommend the following:

1. Existing analysis methods (Generalized Linear Model) are retained for calculating log collated indices (LCI) for individual butterfly species.
    a. Analytical development work is undertaken to extend existing GLMs to incorporate weekly butterfly count data.
2. Multi-species indicators are calculated as a geometric mean of individual species indices, accounting for missing years.
3. Fitting post-hoc, a state-space model (SSM) is used to assess trends in multi-species butterfly indicators.
    a. Analytical development work is undertaken to develop statistical models to fit SSMs with the open source statistics package R to allow models to be extended and to remove the dependence on the proprietary software package TRENDSPOTTER.

References.

Anganuzzi, A.A. (1993) A comparison of tests for detecting trends in abundance indices of dolphins. *Fishery Bulletin*, 91, 183-194.

Besbeas, P., Freeman, S.N. and Morgan, B.J.T (2005) The potential of Integrated Population Modelling. *Australian and New Zealand Journal of Statistics*, 47, 1, 35-48.

Botham, M.S., Brereton, T.M., Middlebrook, I., Cruickshanks, K.L. & Roy, D.B. (2008) United Kingdom Butterfly Monitoring Scheme report for 2007. CEH Wallingford.

Brereton, T.M., Warren, M.S., Roy, D.B. & Stewart, K. (2008). The changing status of the Chalkhill Blue butterfly *Polyommatus coridon* in the UK: the impacts of conservation policies and environmental factors. *Journal of Insect Conservation*, 12 (6): 629-638

Brooks, S.P., Freeman, S.N., Greenwood, J.J.D., King, R. and Mazzetta, C. (2008) Quantifying conservation concern – Bayesian statistics, birds and the Red Lists. *Biological Conservation*, 141, 5, 1436-1441.

Camp, R.J., Seavy, N.E., Gorresen, P.M. and Reynolds, M.H. (2008) A statistical test to show negligible trend: comment. *Ecology*, 89, 1469-1472.

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

Freeman, S.N., Baillie, S.R. and Gregory, R.D. (2001) Statistical analysis of an indicator of population trends in farmland birds. BTO Research Report No. 251, British Trust for Ornithology, Thetford.

Freeman, S.N., Noble, D.G., Newson, S.E. and Baillie, S.R. (2007) Modelling population changes using data from different surveys: the Common Birds Census and the Breeding Bird Survey. *Bird Study*, 54, 1, 61-72**.**

Gregory, R.D. et al. (2007) Population trends of widespread woodland birds in Europe. *Ibis* 149, 78-97, Suppl. 2.

Gregory, R.D. et al. (2008) The generalization and use of bird population indicators in Europe. *Bird Conservation International* 18, S223-S244 Suppl. 1.

Harvey, A.C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Hastie, T.J. and Tibshirani, R. (1990) *Generalized Additive Models*. Chapman and Hall.

Marchant, J.H., Hudson, R., Carter, S.P. and Whittington, P.A. (1990) *Population Trends in British Breeding Birds*. British Trust for Ornithology, Tring.

Mazzetta, C., Brooks, S.P. and Freeman, S.N. (2007) On smoothing trends in population index modelling. *Biometrics*, 63, 4, 1007-1014.

McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models*. Chapman and Hall.

Nelder, J.A. and Wedderburn, R. (1972) Generalized Linear Models. *Journal of the Royal Statistical Society, Series A* 132, 370-384.

Rothery, P. & Roy, D. B. (2001) Application of generalized additive models to butterfly transect count data. *Journal of Applied Statistics*, 28, 897-909.

Roy, D. B., Rothery, P., Moss, D., Pollard, E. & Thomas, J. A. (2001) Butterfly numbers and weather: predicting historical trends in abundance and the future effects of climate change. *Journal of Animal Ecology*, 70, 201-217.

Siriwardena, G.M., Baillie, S.R., Buckland, S.T., Fewster, R.M., Marchant, J.H. and Wilson, J.D. (1998) Trends in the abundance of farmland birds: a quantitative comparison of smoothed Common Birds Census indices. *Journal of Applied Ecology*, 35, 24-43

Soldaat, L., Visser, H., van Roomen, M. and van Strien, A. (2007) Smoothing and trend detection in waterbird monitoring data using structural time-series analysis and the Kalman filter. *Journal of Ornithology*, 148, S351-S357 Suppl. 2.

ter Braak, C.J.F., van Strien, A.J., Meijer, R. and Verstrael, T.J. (1994) Analysis of monitoring data with many missing values: which method? In Hagemeijer, W. & Verstrael, T. (eds.) *Bird Numbers 1992. Distribution, Monitoring and Ecological Aspects*: 663-673. Proceedings of the 12th International Conference of the International Bird Census Council and European Ornithological Atlas Committee. Beek-Ubbergen, The Netherlands: Statistics Netherlands, Voorburg/Heerlen & SOVON.

Visser, H. (2004) Estimation and detection of flexible trends. *Atmospheric Environment* 38, 4135-4145.

Figure 1a. Smoothed butterfly indicators for ‘wider countryside’ species.

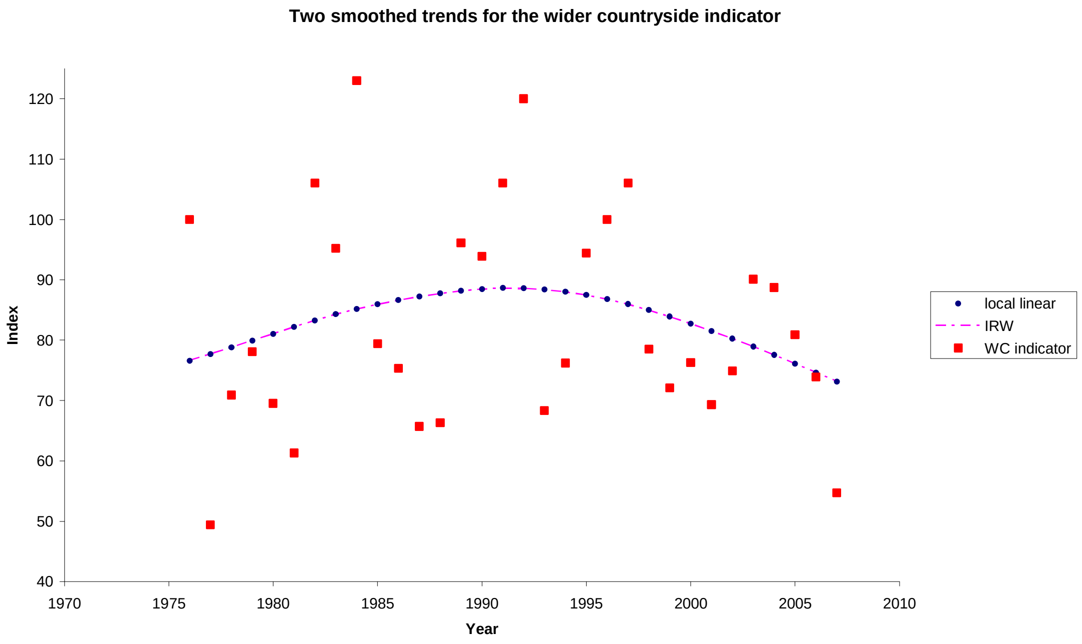

Figure 1b. Standard deviations (SD) of smoothed trends are little altered whether 5 (SD5; broken line) or 10 (SD10; solid line) observations are omitted from the loglikelihood in initialising the filter.