



Effects of choice of baseline on the uncertainty of population and biodiversity indices

Jonas Knapé¹ 

Received: 16 December 2021 / Revised: 20 October 2022 / Accepted: 24 October 2022 /

Published online: 21 November 2022

© The Author(s) 2022

Abstract

Many monitoring programs provide annual indices of relative change over time in some quantitative measure of ecological status, such as population abundance or species richness. These indices are usually scaled relative to a reference year so that they represent change in ecological status compared to this particular year. An issue with this approach is that uncertainty about ecological status in the reference year can propagate into large uncertainty in all other index values. Taking instead the mean of the ecological status over several years as the reference—a reference period—may reduce uncertainty in indices. At present, this approach is not commonly used in practice. I quantitatively evaluate how the choice of reference period affects the uncertainty of two variants of population indices, either estimated independently each year or smoothed over several years, for 100 bird species using monitoring data. Short reference periods containing years early in the series lead to reduced uncertainty in independently estimated index values, but not in smoothed indices, compared to when using a single reference year. When a long reference period was used, uncertainty was substantially reduced for independently estimated annual indices in particular, but also for smoothed indices. An exception to the reduction in uncertainty with the length of the reference period was found when indices are constrained to be log-linear. Given an appropriate model and indices that are not strictly log-linear, using smoothing and/or reference the periods can be useful ways of reducing irrelevant uncertainty in the presentation of indices.

Keywords Baseline · Biodiversity index · Biodiversity trend · Monitoring · Population index · Population trend

Handling Editor: Luiz Duczmal

✉ Jonas Knapé
jonas.knape@slu.se

¹ Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden

1 Introduction

Ecological monitoring programs often produce annual estimates of abundance or biodiversity to assess changes in the status of populations or ecosystems (Marsh and Trenham 2008; Fraixedas et al. 2020). Estimating the total number of individuals in a population, or species in a community, is however a difficult task. For birds and butterflies, for instance, regional population estimates may be derived from transect counts where the proportion of missed individuals cannot be quantified and it may not be clear which spatial area the individuals counted belong to (Ralph et al. 1995, van Swaay et al. 2008). Raw abundance estimates therefore often have no easily interpreted scale. To anchor raw estimates of the ecological quantity of interest, and to put a direct focus on temporal change rather than absolute level, they are usually rescaled relative to some baseline into index values. The use of a baseline gives index values a meaningful scale, and indices are interpreted as the proportional change in the ecological quantity relative to the baseline.

Each choice of baseline gives a different index, and therefore has different associated uncertainty. A standard choice of the baseline is the first year of study (Gregory et al. 2019) so that indices represent the change in the ecological quantity relative to the first year. An issue with this approach is that uncertainty in the raw estimate in the reference year propagates into the uncertainty of all subsequent index values. This may considerably inflate uncertainty (Buckland and Johnston 2017), and is particularly problematic in cases where uncertainty of the raw estimate in the reference year is larger than in subsequent years, which is often the case if the reference year is the first year of the study and sampling effort increases over time. One approach to counter this effect is to select a year with a somewhat lower uncertainty as the reference (Fedy and Aldridge 2011), but this still means that uncertainty in the single reference year propagates into all indices. An alternative is to use the mean over multiple years instead of a single year as the reference to try to get a more stably estimated baseline. This kind of reference is not widely used in practice (but see e.g. Carlson et al. 2012, Knape 2016, Gregory et al. 2019). Another approach that has been suggested to reduce the influence of the choice of baseline (Buckland and Johnston 2017) is to use smoothing methods (Siriwardena et al. 1998; Fewster et al. 2000; Buckland et al. 2005; Soldaat et al. 2017; Harrison et al. 2014; Knape 2016). By smoothing indices over multiple years more stable indices may be obtained, but also in this case using a single reference year to anchor the smoothed index is common practice.

A broader empirical and quantitative assessment of how the choice of reference affects the uncertainty of population indices is currently lacking, and is the aim of this study. I compare indices defined from a single reference year to indices defined from reference periods consisting of sequences of years of varying lengths. I examine how the choice of reference period affect the uncertainty of annual indices estimated independently each year, and of smoothed index estimates obtained from GAMMs, for 100 bird species in Sweden.

2 Methods

For ease of presentation, I discuss abundance indices in the following sections, but the general ideas apply more broadly to ecological indices of temporal change, and particularly to biodiversity indices. Abundance indices are often derived from models with a log-link and, again for the sake of presentation, we assume that we have raw unscaled annual abundance indices over time at the log scale, $\hat{\mu}_1, \hat{\mu}_2, \dots$. These may, for example, be estimated from fixed year effects in a Poisson GLM or, in the case of smoothed indices, from a GAM with a Poisson response. If year 1 is used as the reference, then a standard relative abundance index in year t is

$$\hat{I}_t = \frac{\exp(\hat{\mu}_t)}{\exp(\hat{\mu}_1)} \tag{1}$$

To evaluate the uncertainty of indices we focus on the variance of $\log(\hat{I}_t)$, from which approximate confidence intervals and standard errors can be computed. The reason for focusing on the variance at the log scale is that it is unaffected by simple scaling of the index, e.g. the variance of $\log(\hat{I}_t)$ is identical to that of $\log(100\hat{I}_t)$. The variance of $\log(\hat{I}_t)$ can be expressed as

$$V(\log(\hat{I}_t)) = V(\hat{\mu}_t - \hat{\mu}_1) = V(\hat{\mu}_t) + V(\hat{\mu}_1) - 2\text{Cov}(\hat{\mu}_t, \hat{\mu}_1) \tag{2}$$

This shows that the uncertainty of both $\hat{\mu}_1$ and $\hat{\mu}_t$ contribute to the variance of $\log(\hat{I}_t)$, and that the variance of $\hat{\mu}_1$ will dominate if the uncertainty of $\hat{\mu}_1$ is larger than that of $\hat{\mu}_t$, which is often the case in practice.

The approach evaluated here is to try to reduce the part of the variance in $\log(\hat{I}_t)$ that is due to uncertainty about the index in the reference year by using multiple reference years. Using 2 years as the reference we can define an alternative index

$$\hat{I}_t = \frac{\exp(\hat{\mu}_t)}{(\exp(\hat{\mu}_1) + \exp(\hat{\mu}_2))/2} \tag{3}$$

The arithmetic mean of the raw index in the first 2 years is the reference for this index, and we would often expect the alternative index to have less uncertainty due to the denominator being more precisely estimated. This idea can be extended to use the mean over the first l years as the reference:

$$\hat{I}_t = \frac{\exp(\hat{\mu}_t)}{\frac{1}{l} \sum_{j=1}^l \exp(\hat{\mu}_j)} \tag{4}$$

As l increases we would expect the uncertainty of $\log(\hat{I}_t)$ to decrease further. Whether, and to what extent, this happens in practice is the focus of this paper. I will explore this in an analysis of monitoring data but first briefly describe how the uncertainty of the reference period indices may be computed in practice.

2.1 Computing uncertainty estimates for reference period indices

Compared to single year reference indices, $\log(\hat{I}_t)$ for reference period indices are non-linear as a function of $\hat{\mu}_j$ in the reference period. This can make it more difficult to compute uncertainty estimates. One approach to do so is to use a delta approximation (see Appendix A). A second approach is simulation methods such as bootstrapping where all $\hat{\mu}_t$ are generated according to some simulation procedure, and the non-linear function $\log(\hat{I}_t)$ is computed from the samples. A third approach is Bayesian methods based on Monte Carlo integration, used for example for indices for the North American breeding bird survey (Link and Sauer 2002). In such cases, Monte Carlo samples of the $\hat{\mu}_t$ will usually be available and $\log(\hat{I}_t)$ can be computed for each sample to yield a posterior distribution. A fourth option would be to use the geometric mean of the $\exp(\hat{\mu}_j)$ as the reference instead of the arithmetic mean. Then $\log(\hat{I}_t)$ would be linear as a function of $\hat{\mu}_j$ and the uncertainty could be computed from contrasts.

Reference period indices are currently implemented in the R-packages `rtrim` (Bogaart et al. 2020) and `poptrend` (Knape 2016). The `rtrim` package uses a delta approximation (Appendix A). The `poptrend` package, used in the case study on Swedish birds and simulations below, instead uses a simulation approach. Simulated parameter estimates are drawn from a multivariate normal distribution with covariance matrix equal to a covariance matrix for the parameter estimates (Wood 2006b; Mandel 2013; Harrison et al. 2014).

2.2 Case study

To investigate how multiple reference years affect the uncertainty of population indices in practice, I analyzed data from the Swedish Bird Survey (Lindström and Green 2020). These data consist of annual line transect counts of birds from about 700 survey routes spread across a regular grid over Sweden. Not all routes are surveyed in every year. The survey was initiated in 1996 when 47 routes were surveyed, increasing to 84 in 1997, 166 in 1998, 179 in 1999, and 203 in 2000. The number of routes then continued to increase and 400–500 routes are now surveyed annually. Because so few routes were surveyed in the first 2 years, 1998 is used as the single reference year to compute official indices.

I ranked the species in the survey according to the number of non-zero counts across all years and routes. I then selected the 100 species with the most non-zero counts for analyses of the impact of the baseline on index uncertainty. For each species, I first removed routes that had only zero counts and then fitted (I) a negative binomial regression model with a log link, and route and year as fixed factors and (II) a negative binomial GAMM model with log link, route as a fixed factor, and year both as a smooth function and as a random effect (Knape 2016). The smooth function was implemented as a cubic regression spline. To get a uniform analysis across all 100 species I fixed the number of degrees of freedom in the GAMM at 8. This conforms with other empirical studies of bird census data (Fewster et al. 2000). I fixed the degrees of freedom since model selection of degrees of freedom in the GAMM analysis can lead to smooth functions that are near linear (1 degree of freedom). In this case there is a simple

approximate relationship between uncertainty and the choice of reference period, which is examined further in Appendix B. The random time effect in the GAMM model was included to handle short term variation in abundance, and can affect uncertainty of the estimated trend (Knape 2016). The overdispersion parameter of the negative binomial distribution was estimated rather than fixed a priori, but was not included in the covariance matrix used to estimate uncertainty. Indices were then computed from the estimated year effects for (I) (hereafter referred to as ‘independently estimated annual indices’), or from the estimated smooth function evaluated at the years of interest (see below) for (II) (referred to as ‘smoothed indices’).

To evaluate the effect of choice of reference year or period on uncertainty in the resulting indices, I compared the uncertainty of index values representing year 2020 computed from different reference periods. I used two sets of reference periods, one containing years early in the series when the number of surveyed routes was low, and another containing years in the middle part of the series when more routes were surveyed. Specifically, the first set had reference periods of varying lengths that all ended in 2000 (i.e. 2000, 1999–2000, 1998–2000, etc), and the second set had periods that all ended in 2010 (2010, 2009–2010, 2008–2010 etc). The longest reference period used all 15 years from 1996 to 2010. For the same species and model, all reference period indices were computed from the same model fit. In other words, all years were included in model fits irrespective of the reference period.

For all models and baselines, the amount of uncertainty was measured via the width of a 95% confidence interval for the log scale index in 2020. Data were analyzed using the R-package `poptrend` (Knape 2016), which uses `mgcv` (Wood 2006a) as the model fitting engine, and computes confidence intervals using the simulation procedure described above (Wood 2006b; Mandel 2013).

To complement the case study with a more controlled set up, I also analyzed simulated data. As the results were largely similar to the results of the case study the details are provided in Appendix C. R-code for the analyses is available in Supplement 1.

3 Results

3.1 Independently estimated annual indices

In the case of reference periods in the start of the series (ending in 2000), increasing the number of years in the reference period from one to two reduced the width of confidence intervals by between a few percent and up to almost 60% with a median of 18% (Fig. 1a). The median reduction when using three years was around 24% compared to the single year 2000, while also including the first 2 years with fewer routes sampled on average did not lead to further reductions but instead to slight increases. In a very small number of cases, a reference period including years early in the series led to higher uncertainty than for a single reference year.

Using a single reference year (2010) in the middle of the series gave uncertainty in the same range as when using multiple reference years early in the series (Fig. 1b). With longer reference periods ending in the middle of the series the width of confidence intervals were reduced by up to over 60%, with a median of over 40% and with at least

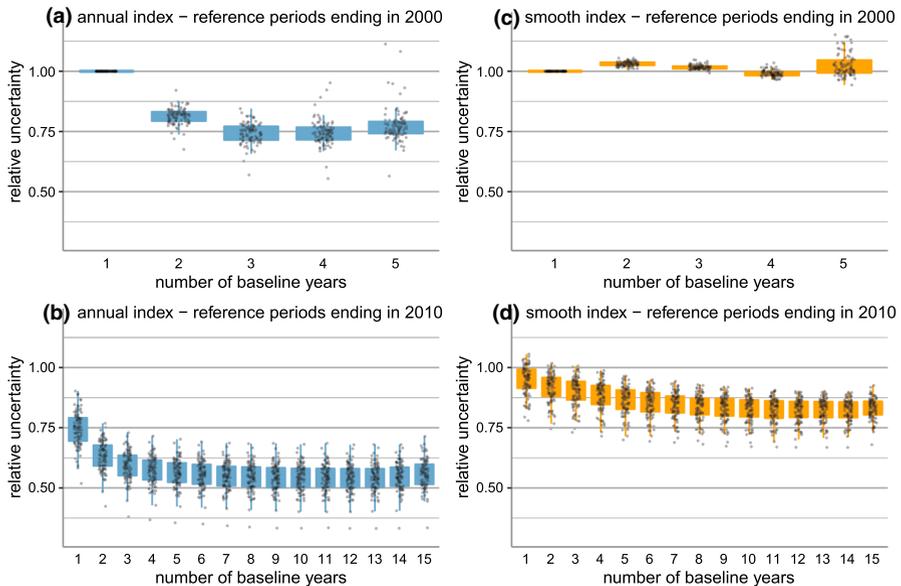


Fig. 1 Boxplots of the relative uncertainty of indices for 100 bird species for independently estimated annual indices (**a** and **b**) and for smoothed indices (**c** and **d**) when using reference periods of varying length ending in 2000 (**a** and **c**) or ending in 2010 (**b** and **d**). Uncertainty is measured via the width of 95% confidence intervals at the log scale for indices in year 2020. In panels **a** and **b**, uncertainty is relative to the uncertainty of independently estimated annual indices with year 2000 as the baseline, and in **c** and **d** relative to smooth indices with 2000 as the baseline

25% for each species, compared to indices with year 2000 as the reference. Including the first few years in these reference periods led to only slight increases in uncertainty (Fig. 1b).

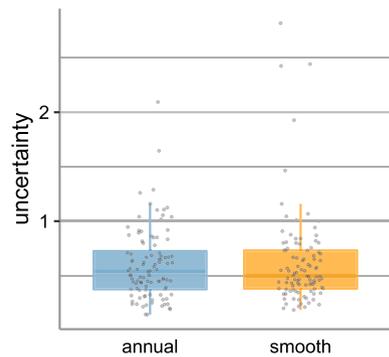
3.2 Smoothed indices

For reference periods in the start of the series, more baseline years did not have a strong effect on uncertainty in the smoothed indices (Fig. 1c). Indices with reference period ending in the middle of the series had lower uncertainty (Fig. 1d). This was not only a consequence of more recent years being included in the reference period as extending the reference period backwards in time gave less index uncertainty than using only year 2010 as the reference (Fig. 1d). This is the opposite of what one would expect for a log-linear index for which uncertainty is mainly determined by the location of the midpoint of the reference period (Appendix B).

The magnitude of reduction in uncertainty for longer reference periods was lower than for independently estimated annual indices. The maximum reduction in uncertainty was on average around 15% compared to using year 2000 as the reference (Fig. 1d).

All results reported above are scaled by the width of confidence intervals for an index with the single year 2000 as the reference period ((width of the confidence

Fig. 2 Boxplots of the uncertainty of indices for year 2020 for 100 bird species and independently estimated annual indices and smooth indices when using the single reference year 2000. Uncertainty is measured via the width of 95% confidence intervals at the log scale for indices in year 2020



interval for index with reference period)/(width of the confidence interval for index with year 2000 as the reference)). This scaling removes heterogeneity in estimation errors due to species properties, such as abundance and overdispersion. The variability in this scaling factor (the denominator of the ratio) is presented in Fig. 2.

An illustration of the different models fitted to data on willow warblers can be found in Appendix D (Fig. 8).

4 Discussion

Choosing a baseline to anchor indices of population or biodiversity status is often necessary for meaningful presentation of change. The typical baseline choice is a single year in early parts of the series. The results of this study show that when annual indices are independently estimated, uncertainty can be greatly reduced by redefining the index using a longer reference period as the baseline. A single reference year may give the impression that uncertainty about population change or biodiversity is large, while in fact the main portion of uncertainty comes from asserting the level in the specific reference year. A longer reference period may therefore be beneficial and give a more accurate picture of uncertainty.

Previous studies have suggested that smoothed indices are less sensitive to the choice of reference year or period (Buckland and Johnston 2017). The results of this study confirm this empirically in that the length of reference periods had less impact on uncertainty than for independently estimated annual indices, and for all baselines there was less variation in uncertainty. Even so, smooth indices with long reference periods had, on average, around 15% less uncertainty than smooth indices with a single reference year. Longer reference periods therefore can be useful also for smooth indices, as long as the smooth index is not near linear in which case there is not much to gain from using reference periods instead of a single year (Appendix B).

When deciding on a baseline, the first priority should be a choice that reflects the purpose of the index. If the purpose is to compare the current status to the status in a specific year in the past, then a single reference year is appropriate despite potentially high uncertainty. In other cases, a reference period may fit well with the purpose of the index. Using a moving ten year period as the reference may for example fit well

with IUCN red list assessments where change during the last 10 year period is an important criterion, or in the context of consequences of climate change a reference period coinciding with a climate normal period such as 1961–1990 could be suitable. Often there may not be an obvious year or period against which comparisons should be made, as the main purpose of many indices is in understanding how the size of a population, or the biodiversity of a community, has changed relative to the past in a more loose sense. In such cases choosing a baseline so that the index does not convey irrelevant uncertainty should be an important consideration. In light of the results here, using the mean over a large part, or all, of the series, or over the previous ten-year period (Buckland and Johnston 2017) seem like reasonable default choices in such situations.

Alternative suggestions have been to entirely remove the influence of baseline years by focusing on the slope or curvature of estimated smooth index curves (Buckland and Johnston 2017). Specifically, p-values for whether the first or second derivative of the smooth curve deviates from zero may be computed (Fewster et al. 2000), and one may use these to produce indicators for periods where the change in the curve, or the slope of the curve, is significant. Such indicators are a highly useful complement to indices, but address a partially different question. They are indicators for periods of change in status, but do not provide direct estimates of the cumulative magnitude of the change. Both of these are of prime interest for monitoring, and can be simultaneously presented in displays of indicators.

Uncertainty is an important but sometimes neglected component of population indices (Fraixedas et al. 2020). It is imperative that uncertainty estimates account for the main sources of error, which mainly comes down to a sound choice of model for producing raw index estimates. Given that important sources of error in the data have been properly accounted for, indices should be presented in a way that does not include irrelevant uncertainty. Smoothing indices and/or using longer reference periods, are useful approaches to achieve this.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10651-022-00550-7>.

Acknowledgements I am grateful to Andreas Lindén for comments and discussion. This study was funded by grant 2017-1064 from the Swedish Research Council FORMAS.

Funding Open access funding provided by Swedish University of Agricultural Sciences.

Declarations

Conflict of interest The author has no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted

by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: Delta approximations

Delta approximation for reference period indices at the log scale

For a scalar function g taking input in the form of a vector x , the delta approximation for the variance of $g(X)$ where X is multivariate random variable with mean vector μ and variance matrix Σ is Ver Hoef (2012)

$$V(g(X)) \approx \nabla g(\mu)^T \Sigma \nabla g(\mu) \tag{A.1}$$

We first assume that we have estimates of abundance or biodiversity at the log scale, μ_t , as in the main text, and also that we have estimates of uncertainty for these in the form of variances $V(\mu_t)$ and covariances $\text{Cov}(\mu_j, \mu_k)$. The delta approximation gives us a way of approximating the variance of $\log(I_t)$ from this information about the μ_t .

Here, we are interested in the function $\log(I_t) = \mu_t - \log(1/m \sum_{j=1}^m e^{\mu_j})$. We restrict attention to the case $t > m$, minor adjustments below would be needed to cover the case $t \leq m$. The vector $\nabla \log(I_t)$ contains the derivatives of $\log(I_t)$ with respect to μ_j for $j = 1, \dots, m - 1, m, t$. For $j \leq m$ the derivative is

$$\frac{\partial \log(I_t)}{\partial \mu_j} = \frac{-e^{\mu_j}}{\sum_{l=1}^m e^{\mu_l}}$$

and for t the derivative is equal to 1. The gradient therefore becomes:

$$\nabla \log(I_t) = \begin{bmatrix} \frac{\partial \log(I_t)}{\partial \mu_1} \\ \vdots \\ \frac{\partial \log(I_t)}{\partial \mu_m} \\ \frac{\partial \log(I_t)}{\partial \mu_t} \end{bmatrix} = \begin{bmatrix} \frac{-e^{\mu_1}}{Z} \\ \vdots \\ \frac{-e^{\mu_m}}{Z} \\ 1 \end{bmatrix}$$

where $Z = \sum_{l=1}^m e^{\mu_l}$. The variance matrix Σ is defined from the variance and covariance terms $\text{Cov}(\mu_j, \mu_k)$. To compute the approximate variance of $\log(I_t)$ in practice it is often convenient to use (A.1) directly by plugging in the gradient and the covariance matrix. However, we can also expand the matrix product to arrive at a direct expression for the variance as shown below.

The $2m + 1$ terms involving μ_t in the sum $\nabla \log(I_t)^T \Sigma \nabla \log(I_t)$ combine into:

$$\begin{aligned} & V(\mu_t) \left(\frac{\partial \log(I_t)}{\partial \mu_t} \right)^2 + 2 \sum_{j=1}^m \text{Cov}(\mu_t, \mu_j) \frac{\partial \log(I_t)}{\partial \mu_t} \frac{\partial \log(I_t)}{\partial \mu_j} \\ &= V(\mu_t) - 2 \sum_{j=1}^m \text{Cov}(\mu_t, \mu_j) \frac{e^{\mu_j}}{Z}. \end{aligned}$$

The m^2 terms not involving μ_t combine into:

$$\begin{aligned} & \sum_{j=1}^m V(\mu_j) \left(\frac{\partial \log(I_t)}{\partial \mu_j} \right)^2 + \sum_{j \neq k} \text{Cov}(\mu_j, \mu_k) \frac{\partial \log(I_t)}{\partial \mu_j} \frac{\partial \log(I_t)}{\partial \mu_k} \\ &= \sum_{j=1}^m V(\mu_j) \frac{e^{2\mu_j}}{Z^2} + 2 \sum_{j=1}^m \sum_{k=j+1}^m \text{Cov}(\mu_j, \mu_k) \frac{e^{\mu_j + \mu_k}}{Z^2} \end{aligned}$$

Adding the two sums together gives

$$\begin{aligned} V(\log(I_t)) \approx & V(\mu_t) - 2 \sum_{j=1}^m \text{Cov}(\mu_t, \mu_j) \frac{e^{\mu_j}}{\sum_{l=1}^m e^{\mu_l}} + \sum_{j=1}^m V(\mu_j) \frac{e^{2\mu_j}}{(\sum_{l=1}^m e^{\mu_l})^2} \\ & + 2 \sum_{j=1}^m \sum_{k=j+1}^m \text{Cov}(\mu_j, \mu_k) \frac{e^{\mu_j + \mu_k}}{(\sum_{l=1}^m e^{\mu_l})^2}. \end{aligned}$$

Delta approximation at the arithmetic scale

The delta approximation can also be used on the arithmetic scale to compute standard errors of \bar{I}_t from the variance of $M_j := e^{\mu_j}$. This is what the rtrim package does to compute the uncertainty for reference period indices. In this case the gradient is

$$\nabla I_t = \begin{bmatrix} \frac{\partial I_t}{\partial M_1} \\ \vdots \\ \frac{\partial I_t}{\partial M_m} \end{bmatrix} = \begin{bmatrix} \frac{-m M_t}{(\sum_{l=1}^m M_l)^2} \\ \vdots \\ \frac{-m M_t}{(\sum_{l=1}^m M_l)^2} \\ \frac{m}{\sum_{l=1}^m M_l} \end{bmatrix}$$

and Σ is the covariance matrix for the M_j .

Appendix B: Uncertainty of linear indices

We here examine properties of indices that are strictly linear at the log-scale. When the trend line is assumed to be log-linear a simple approximate expression can be derived for the uncertainty of indices as a function of the reference period. One can then show that the relative uncertainty of the linear indices does not depend on the details of the sampling model.

Assume that uncertainty is evaluated for year $t0 + n$ (say 2020 = 2010 + 10 in our case) and that the reference period is (in R-like notation) $t0 + (m - l + 1) : m$ so that m is the last year of the reference period (relative to $t0$) and l is the number of years in the reference period. To simplify notation, assume that $t0$ corresponds to year 2010 and set $t0 = 0$, which is equivalent to redefining t as the number of years since 2010. The reference period then is simply $(m - l + 1) : m$.

To derive an approximate estimate of the uncertainty of the index in year n we compute the variance of the index

$$\hat{\mu}_n - 1/l \sum_{k=m-l+1}^l \hat{\mu}_k.$$

Recall that $\hat{\mu}$ was defined at the log scale so that we are effectively using the geometric mean over the reference period rather than the arithmetic mean as in the case study. This is an approximation of the uncertainty for indices defined from arithmetic mean reference periods, but it holds approximately if the annual indices do not vary considerably, i.e. if the slope of the trend line is not steep. It turns out to work quite well for linear estimates of trends for the 100 species in the case study (Fig. 3).

If the index is derived from a linear model we have

$$\hat{\mu}_t = \hat{\alpha} + \hat{\beta}t$$

for some estimated intercept $\hat{\alpha}$ and slope $\hat{\beta}$. The (log-scale) index in year n is then

$$\hat{\mu}_n - 1/l \sum_{k=m-l+1}^l \hat{\mu}_k = \hat{\beta}(n - 1/l \sum k) = \hat{\beta}(n - m + (l - 1)/2)$$

and its variance is

$$V\left(\hat{\mu}_n - 1/l \sum_{k=m-l+1}^l \hat{\mu}_k\right) = V(\hat{\beta})(n - m + (l - 1)/2)^2.$$

Uncertainty is measured via the width of confidence intervals for the index in year n , which should be approximately proportional to the standard deviation of the index. The uncertainty of the index with reference period $(m - l + 1) : m$ relative to the index with year 0 (corresponding to year 2010) as the reference is then approximately

$$\frac{SD(\hat{\mu}_n - 1/l \sum_{k=m-l+1}^l \hat{\mu}_k)}{SD(\hat{\mu}_n - \hat{\mu}_0)} = \frac{n - m + (l - 1)/2}{n} = 1 - \frac{1}{n}m + \frac{1}{n} \frac{l - 1}{2}$$

This expression suggests that the reduction in uncertainty for linear indices is mainly determined by the proximity of the center (midpoint) of the reference period to the evaluation year. Note that the variance of the slope estimate, which depends on the details of the sampling model, cancels out of the relative uncertainty calculated above, so this result is independent of the sampling model.

To check how linear indices behave for the case study of 100 bird species, I fitted models with a linear effect of year plus random year and site effects at the log-scale and a negative binomial response distribution (analogously to the setup in the main text). Relative uncertainty is shown in Fig. 3, and corresponding scaling factors in Fig. 4. Extending reference periods backwards in time (i.e. keeping m fixed but increasing l)

Fig. 3 Boxplots of the relative uncertainty of indices for 100 bird species for log-linear indices when using reference periods of varying length ending in 2000 (top panel) or ending in 2010 (bottom panel). Uncertainty is measured via the width of 95% confidence intervals at the log scale for indices in year 2020, and is relative to the uncertainty of log-linear indices with year 2000 as the baseline. Green dots show the expected relative uncertainty computed from the linear approximation

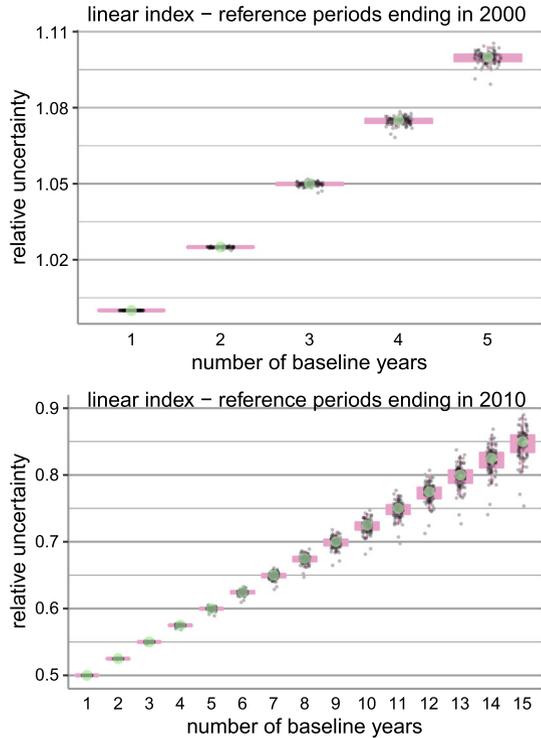
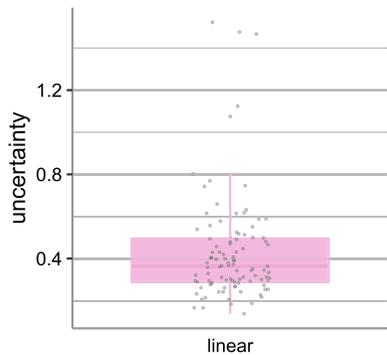


Fig. 4 Boxplot of the uncertainty of log-linear indices for year 2020 for 100 bird species when using the single reference year 2000. Uncertainty is measured via the width of 95% confidence intervals at the log scale for indices in year 2020



leads to increased uncertainty (Fig. 3), which may seem counter intuitive. However, this is due to the resulting backward shift of the center of the reference period. If we alternatively extend reference periods both forward and backward in time while keeping their centers fixed, the uncertainty of linear indices is largely unaffected by the length or reference period. This is shown in Fig. 5 where reference periods are centered at year 2003 and simultaneously extended both forward and backward in time (2003, 2002–2004, 2001–2005 etc.).

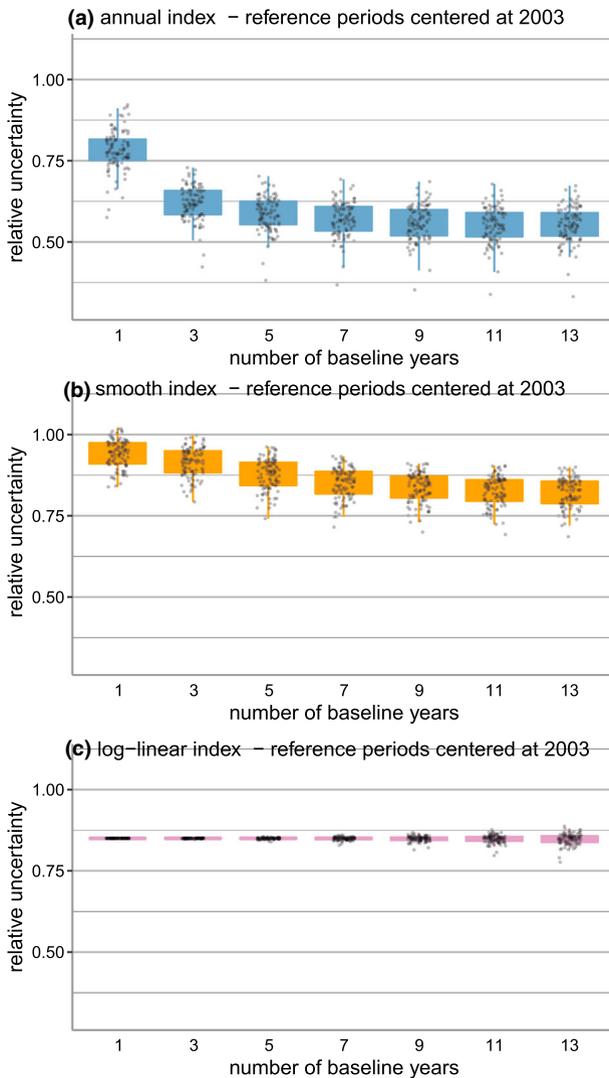


Fig. 5 Boxplots of the relative uncertainty of indices for 100 bird species for independently estimated annual indices (a), smoothed indices (b), and log-linear indices (c), when using reference periods of varying length centered at 2003. Uncertainty is measured via the width of 95% confidence intervals at the log scale for indices in year 2020, and is relative to the uncertainty of, respectively, annual, smooth and log-linear indices with year 2000 as the baseline

Appendix C: Simulations

To check that results from the case study are not simply due to properties of the data (such as unbalanced and missing data, or deviations from parametric assumptions) I additionally analyzed simulated data. The simulation set-up mimicked the case study

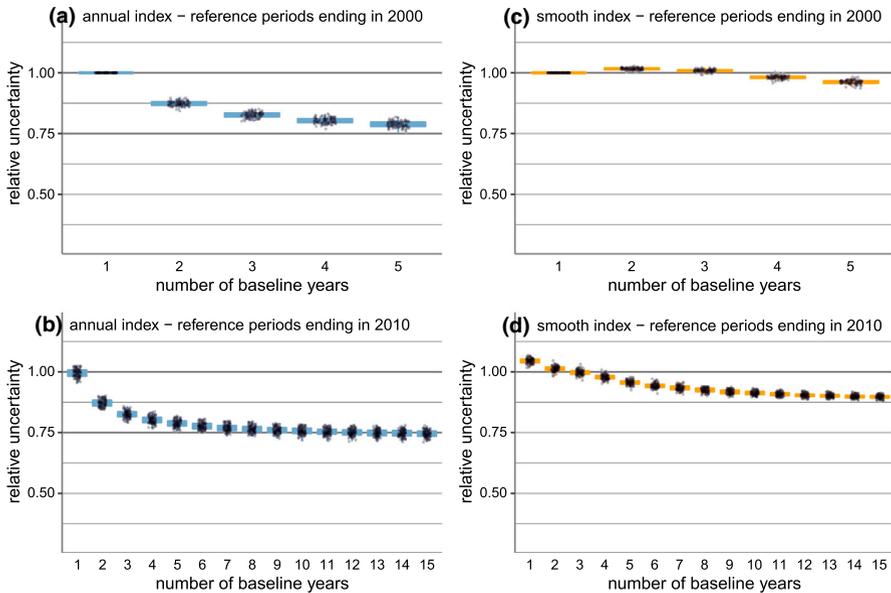
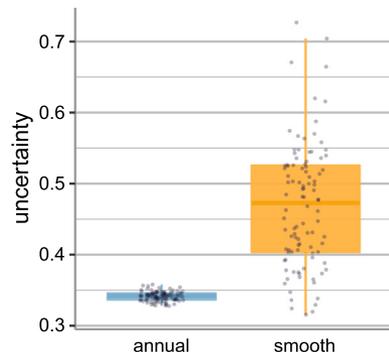


Fig. 6 Boxplots showing the relative uncertainty of indices for 100 simulated data sets for independently estimated annual indices (**a** and **b**) and for smoothed indices (**c** and **d**) when using reference periods of varying length ending in 2000 (**a** and **c**) or ending in 2010 (**b** and **d**). Uncertainty is measured via the width of 95% confidence intervals at the log scale for indices in year 2020. In panels **a** and **b** uncertainty is relative to the uncertainty of independently estimated annual indices using year 2000 as the baseline, and in **c** and **d** relative to smooth indices using 2000 as the baseline

with 100 species, 400 sites and for a time period corresponding to 1996–2020, but with no missing data for any of the site and year combinations. I used an overall intercept of 1 at the log scale and site effects were drawn iid from a normal distribution with mean zero and standard deviation 0.5. Year effects were composed of two parts, a quadratic function peaking in 2006, $-0.002(\text{year} - 2006)^2$, and added to that random iid draws from a normal distribution with mean zero and standard deviation 0.1. The negative binomial size parameter was set to 1 in the simulations. The same models as in the main text were fitted to each simulated data set.

Fig. 7 Boxplots of the uncertainty of indices for 100 simulated data sets and independently estimated annual and smooth indices when using the single reference year 2000. Uncertainty is measured via the width of 95% confidence intervals at the log scale for indices in year 2020



Results showed similar patterns in reduction of uncertainty with increasing length of the reference period as for the case study in the main text (Fig. 6, corresponding scaling factors in Fig. 7). However, as the data were simulated with identical sample sizes for all years, independently estimated annual indices had similar uncertainty when the reference period ended in 2000 as when they ended in 2010, and smooth indices with the single reference year 2010 had slightly higher uncertainty than indices with the single reference year 2000.

Appendix D: Example indices

Examples of the different models for data on willow warbler are shown in Fig. 8.

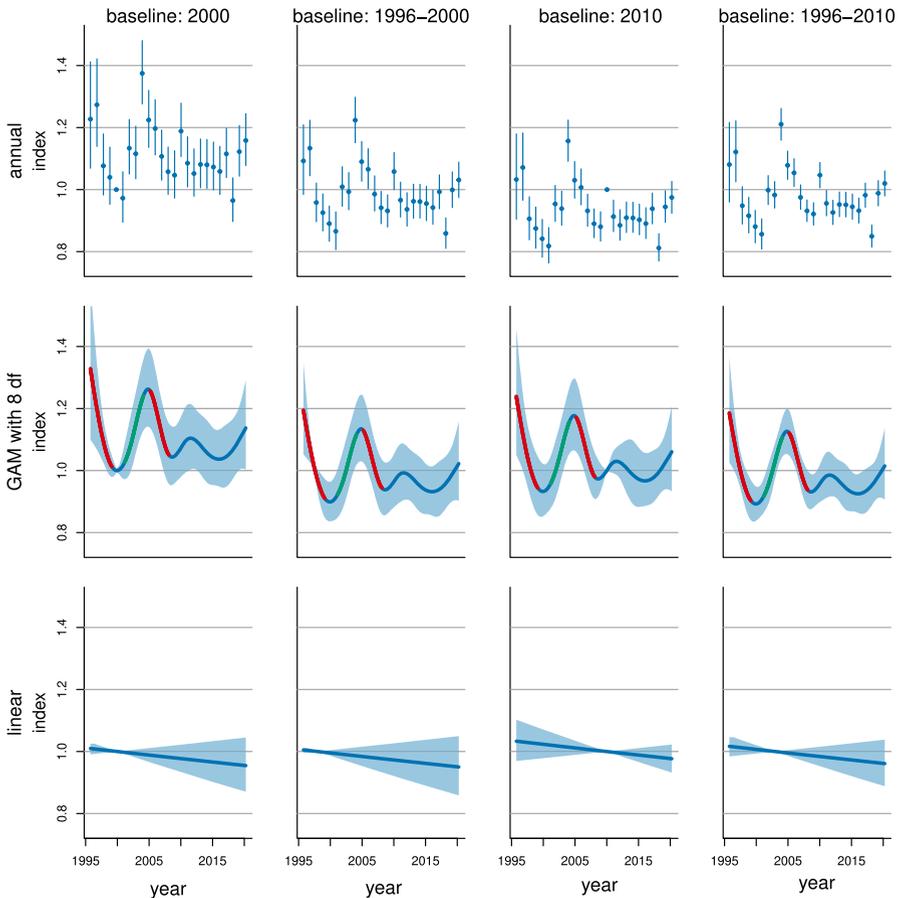


Fig. 8 Comparison of indices and their uncertainty (95% confidence intervals) for the willow warbler using a single reference year early in the series (2000), a reference period with the first five years of the study, a single reference year in the middle part of the series (2010), and a long reference period consisting of all years between 1996 and 2010. The first row shows independently estimated annual indices, the second row shows smooth GAMM indices with 8 degrees of freedom, and the bottom row shows log-linear indices. Green segments show periods of significant increase and red segments significant decrease, computed using finite differences (Fewster et al. 2000)

References

- Bogaart P, van der Loo M, Pannekoek J (2020) rtrim: trends and indices for monitoring data. <https://CRAN.R-project.org/package=rtrim>. Accessed 14 Nov 2020
- Buckland ST, Johnston A (2017) Monitoring the biodiversity of regions: key principles and possible pitfalls. *Biol Conserv* 214:23–34. <https://doi.org/10.1016/j.biocon.2017.07.034>
- Buckland ST et al (2005) Monitoring change in biodiversity through composite indices. *Philos Trans R Soc B* 360:243–254. <https://doi.org/10.1098/rstb.2004.1589>
- Carlson JK et al (2012) Relative abundance and size of coastal sharks derived from commercial shark longline catch and effort data. *J Fish Biol* 80:1749–1764. <https://doi.org/10.1111/j.1095-8649.2011.03193.x>
- Fedy BC, Aldridge CL (2011) The importance of within-year repeated counts and the influence of scale on long-term monitoring of sage-grouse. *J Wildl Manag* 75:1022–1033. <https://doi.org/10.1002/jwmg.155>
- Fewster RM et al (2000) Analysis of population trends for farmland birds using generalized additive models. *Ecology* 81:1970–1984
- Fraixedas S et al (2020) A state-of-the-art review on birds as indicators of biodiversity: advances, challenges, and future directions. *Ecol Indic* 118:106–128. <https://doi.org/10.1016/j.ecolind.2020.106728>
- Gregory RD et al (2019) An analysis of trends, uncertainty and species selection shows contrasting trends of widespread forest and farmland birds in Europe. *Ecol Indic* 103:676–687. <https://doi.org/10.1016/j.ecolind.2019.04.064>
- Harrison PJ et al (2014) Assessing trends in biodiversity over space and time using the example of British breeding birds. *J Appl Ecol* 51:1650–1660. <https://doi.org/10.1111/1365-2664.12316>
- Knape J (2016) Decomposing trends in Swedish bird populations using generalized additive mixed models. *J Appl Ecol* 53:1852–1861. <https://doi.org/10.1111/1365-2664.12720>
- Lindström Å, Green M (2020) Swedish bird survey: fixed routes (Standardrutterna). <https://doi.org/10.15468/hd6w0r>
- Mandel M (2013) Simulation-based confidence intervals for functions with complicated derivatives. *Am Stat* 67(2):76–81. <https://doi.org/10.1080/00031305.2013.783880>
- Marsh DM, Trenham PC (2008) Current trends in plant and animal population monitoring. *Conserv Biol* 22:647–655. <https://doi.org/10.1111/j.1523-1739.2008.00927.x>
- Ralph CJ, Sauer JR, Droege S (1995) Monitoring bird populations by point counts. In: Gen. Tech. Rep. PSW-GTR-149. Albany: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station. 187 p 149. <https://doi.org/10.2737/PSW-GTR-149>
- Siriwardena GM et al (1998) Trends in the abundance of farmland birds: a quantitative comparison of smoothed Common Birds Census indices. *J Appl Ecol* 35:24–43. <https://doi.org/10.1046/j.1365-2664.1998.00275.x>
- Soldaat LL et al (2017) A Monte Carlo method to account for sampling error in multi-species indicators. *Ecol Indic* 81:340–347. <https://doi.org/10.1016/j.ecolind.2017.05.033>
- van Swaay CAM et al (2008) Butterfly monitoring in Europe: methods, applications and perspectives. *Biodivers Conserv* 17:3455–3469. <https://doi.org/10.1007/s10531-008-9491-4>
- Ver Hoef JM (2012) Who invented the delta method? *Am Stat* 66:124–127. <https://doi.org/10.1080/00031305.2012.687494>
- Wood SN (2006a) Generalized additive models: an introduction with R. CRC Press. <https://doi.org/10.1201/9781420010404>
- Wood SN (2006b) On confidence intervals for generalized additive models based on penalized regression splines. *Aust N Zeal J Stat* 48(4):445–464. <https://doi.org/10.1111/j.1467-842X.2006.00450.x>