Monitoring the biodiversity of regions: key principles and possible pitfalls

S.T. Buckland*
Centre for Research into Ecological and Environmental Modelling, University of St Andrews
The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ, UK

A. Johnston
British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU, UK
Now at: Cornell Lab of Ornithology, 159 Sapsucker Woods Road, Cornell University, NY14850, USA
and Conservation Science Group, Dept of Zoology, University of Cambridge, CB2 3EJ, UK

Abstract

Through the Convention on Biological Diversity (CBD) 2010 and 2020 biodiversity targets, nations committed to reducing the rate of loss of biodiversity. This requires calculating the biodiversity trends in nations, whereas previously, most academic research on quantifying biodiversity concerned communities within relatively small sites. We consider design and analysis issues that CBD targets raise and explore the potential pitfalls for managers of monitoring schemes when statistical principles yield to practical constraints. We list five main criteria that well-designed monitoring programmes should meet: representative sampling locations, sufficient sample size, sufficient detections of target species, a representative sample of species, and a sound temporal sampling scheme. We examine the implications of biodiversity assessments that fail to meet these criteria and suggest ways to alleviate these implications through analytical approaches. We discuss the remarkable potential for wide-scale biodiversity monitoring offered by technological advances and by the rise of citizen science.

Keywords: animal abundance estimation; biased sample; biodiversity trends; Convention for Biological Diversity targets; geometric mean; representative sample

*Corresponding author (steve@st-andrews.ac.uk)

1. Introduction

The 2010 Biodiversity Target of the Convention on Biological Diversity (CBD), set in 2002, had far-reaching consequences for how biodiversity is measured (Butchart et al., 2010). It was superseded by 20 targets for 2020, which have an overall mission to "take effective and urgent action to halt the loss of biodiversity" (CBD, 2011). Thus, long-term biodiversity monitoring programmes are needed, together with effective measures of biodiversity trends, to assess success or failure in meeting the targets (Pereira and Cooper, 2006; Mace and Baillie, 2007; Magurran et al., 2010). Because targets are agreed by nations, it is necessary to measure the biodiversity of nations; that is, we need programmes that allow quantification of biodiversity trends across large geographic regions (Buckland et al., 2011, 2012a, in press). Rodrigues et al. (2014) estimate that most loss of global biodiversity is concentrated in just eight countries (Australia, China, Colombia, Ecuador, Indonesia, Malaysia, Mexico, and the United States), which highlights the need for effective monitoring by nation.

Ideally, robust and long-term monitoring programmes would enable assessment of changes of biodiversity within countries or large regions. However, many monitoring programmes are targeted towards small spatial areas, or have other drawbacks such as no clear monitoring target, low power to detect change, or biased selection of sites or species (Yoccoz et al., 2001; Peireira and Cooper, 2006; Legg and Nagy, 2006). Although there are many books and articles with guidelines for statistical principles of sampling (e.g. Sutherland, 1996; Manly and Navarro Alberto, 2014), there are various reasons why these principles are often not applied in ecological surveys of nations or large regions. Firstly, several long-term monitoring programmes were established many years ago when principles of survey design were less well-established and when technology and funding landscapes were very different. Ecological inferences from these long-term schemes may be limited by the precision achieved at the start of the survey, even if sample size has subsequently expanded. Secondly, the financial resources, number of surveyors or technology may limit robust inference to a small region, or low power over a large region (Legg and Nagy, 2006; Taylor et al., 2006). Lastly, samples are often spatially or temporally biased, perhaps due to using citizen scientists or the expense of surveying certain areas (Stolar and Nielsen, 2015). All these differences between the ideal statistical sampling protocol and the realised sampling scheme can cause problems when using these data to infer change in biodiversity across a wide region.

In this paper we outline the ideal requirements for large-scale monitoring programmes and discuss the implications for estimates of biodiversity when these are not met. We reference some example surveys that meet criteria for robust ecological inference and some surveys that do not. We discuss the trade-offs between inference from sub-optimal sampling regimes that can be applied widely and inference from ideal sampling regimes that may be restricted to a very few regions or species. We also discuss the conservation implications of sub-optimal sampling regimes to estimate trends in biodiversity.

2. Five criteria for effective monitoring programmes

To estimate biodiversity across broad spatial extents, monitoring programmes are needed that allow temporal trends of multiple species to be estimated for large regions. Well-designed monitoring programmes should meet the following criteria: 1) representative sampling locations, 2) sufficient sample size, 3) sufficient detections of target species, 4) representative sample of species (or all species), 5) a temporal sampling scheme designed to aid valid inference. To assess whether a particular scheme meets these criteria, it is important to have clear monitoring goals. These will include specification of the region, species and timescale that a scheme is designed to monitor.

Firstly, representative sampling locations are needed to ensure that the estimated trends in biodiversity are representative of the region of interest and not biased towards particular habitats or locations. Representative estimates can be achieved in two ways: design-based or model-based. Design-based representativeness requires the sampling locations to be representative and this is often achieved by simple random or stratified random site selection (Buckland et al., 2012a). Model-based representativeness corrects for sampling locations that are not representative by reweighting the contribution of each sample, such that the contribution of samples to the overall trend estimate are representative. For example, reweighting can account for habitats that are sampled in different proportions to the total environment (van Swaay et al., 2008) or countries that contain different proportions of an overall population (Gregory et al., 2005). When a randomized sampling scheme (whether stratified or not) is not feasible, non-representative sampling locations are chosen either by design (for example to target a rare species or an accessible locale) or implicitly (for example by the accumulated decisions of many individual citizen scientists). This non-representative sample generally results in false inferences, because we cannot assume that the sampling location is chosen independently of the trend at that location. However, if care is taken in the selection of sites, then it may be possible to develop model-based analysis methods that account for bias in the sampling.

Secondly, sufficient sample size is required to estimate biodiversity trends with a reasonable precision. If too few sites are sampled, estimates of biodiversity trends will be imprecise, and estimates of the precision may be poor (Carlson and Schmiegelow, 2002; Nielsen et al., 2009). In order to detect changes in the rate of change of biodiversity or a cessation of biodiversity loss, monitoring programmes need to estimate trends with high precision and low bias.

The third criterion for monitoring programmes is that they require sufficient detections of target species. Assuming there are sufficient geographical samples (criterion 2), the number of detections for a given species may be low because it is rare, or because it has low detectability. While it may be more cost-effective to implement a single survey for all species in the community of interest, it may be necessary to have separate schemes for key species, for example to ensure that the range of a rare and restricted species is adequately sampled, or to allow different field methods for those species whose individuals have low detectability under the standard protocol. Ideally, analytical methods will estimate detectability, for example using distance sampling methods (Buckland et al., 2015), double-observer methods (Nichols et al., 2000) or repeat visits and occupancy modelling (MacKenzie et al., 2006).

Fourthly, those species monitored should be representative of all species in the community of interest. Ideally, all species in the target group would be monitored, but if this is not feasible, then careful consideration should be given to selecting species for monitoring; if only common and easily detectable species are monitored, we can have no confidence that biodiversity trend estimates reflect trends in the wider community of interest.

Finally, careful consideration should be given to the temporal element of the survey design. The ideal design might be annual surveys, conducted at the same time each year. A time of year should be selected when rapid change, for example due to migration or appearance of young, is unlikely. For example, songbird numbers in temperate regions tend to be stable early in the breeding season, when males are holding territories and young have not yet fledged. In this case, precision of a given annual estimate is largely a result of sampling variance, and not of short-term population changes, and thus trends are estimated with higher precision. The possibility of phenological changes should be considered, as there may be a trend towards earlier breeding as a result of climate change. Sampling at a fixed time each year may estimate a declining or increasing trend, due to change in time of migration or breeding (Dennis et al., 2013). If it is not possible to survey all sampled locations annually, then a rolling survey might be adopted in which every site is surveyed say every three years, with a third of the sites surveyed each year. Another option is to do a complete survey every few years and there are various other options for temporally unbiased survey designs, such as rotating panel designs, in which a proportion of sites is retained from the previous year (McDonald, 2003). See Duncan and Kalton (1987) and Binder and Hidiroglou (1988) for reviews, and Underwood (2012) for a proposed framework for adapting survey design through time.

Well-designed monitoring programmes will have a clear target ecological community and monitoring region, so that the criteria can be assessed against these targets. Monitoring objectives generally fall into two categories: to describe the trend or explain the trend (or both). Here we focus on surveys that are designed to describe the trend in biodiversity as this directly relates to the CBD targets. Schemes to estimate the drivers of trends or other explanations will have different optimal survey designs, although the principles outlined here will be similar (Hirzel and Guisan, 2002; Maggini et al., 2002).

Once clear objectives have been defined, simulations can be used to assess the ability of a proposed programme to meet the stated objectives. In traditional power analyses, the power is the proportion of simulations that correctly identify a significant trend for a species or community (correctly reject the null hypothesis). Outside of a hypothesis testing framework, simulations can also be used to compare the estimated trends and precision for a range of different survey designs. For example, the proportion of simulated surveys that identify a trend with a given precision (related to the monitoring objectives) might be compared across designs. These comparisons can also account for constraints such as finances or number of surveyors (e.g. Teilmann et al., 2010; Field et al., 2005; Sanderlin et al., 2014). However, we caution that it can be challenging to simulate a realistic community of species across a landscape and overly-simplified simulations of a community may give a false impression of the power of a particular design. Therefore we advocate the use of simulations for comparing survey designs rather than assessing the power directly.

The UK Breeding Bird Survey (BBS, Newson et al., 2005) is an example of a survey that comes closer than most to meeting the above criteria. It aims to survey common breeding birds in the UK, and is based on a stratified random sample of 1km squares, where strata correspond approximately to administrative regions and the sampling intensity in each stratum is in proportion to the number of available observers. This is a design-based sampling strategy and the variation in sampling intensity across strata is accounted for using weights in the analysis; thus the sample is representative and the survey satisfies criterion 1. The survey protocol is that in each sampled square, two transects, each 1km long, are walked, and each detected bird is assigned to one of four categories denoting the distance from the transect: 0-25m, 25-100m, >100m and flying. Approximately 3000 squares are now surveyed twice during each breeding season. The sample size enables reasonably precise annual estimates of population trend for approximately 100 species, or 40% of the UK's breeding birds. The protocol is designed for birds that vocalise or are visible during daylight hours. It is not well-suited to nocturnal species, those that are hard to detect, or those which have a very restricted range (sample size becomes low as there are too few sites that detect the species). The standard trend analyses assume that within any given species, detectability is constant over time so that the counts can be considered to be relative abundance estimates, which was found to be a reasonable assumption for the majority of species (Newson et al., 2013). A further compromise is that the nominal transect line cannot always be followed, and so there tends to be a bias towards placing transects along edge habitats, especially in areas of arable farming, where observers cannot walk through crops. The UK BBS began in 1994 as a replacement for the Common Bird Census (CBC). From 1994–2000, both schemes were run in parallel, allowing calibration of the estimates from the two surveys (Freeman et al., 2007). Together these annual surveys provide estimates of breeding bird population trends from the 1960s to the present day, providing a temporally rich dataset (criterion 5). The CBC was replaced because CBC sites were not selected according to a randomized scheme, and precise trend estimates were restricted to southern Britain, as there were too few CBC sites in the north.

3. Widening the scope of biodiversity monitoring through technological advances

The UK BBS uses knowledgeable birdwatchers as surveyors and it is an example of a citizen science project that is well-designed to allow high-quality inference on species and biodiversity trends. By contrast, many citizen science monitoring projects generate large sample sizes (aiding criterion 2), but have poor representativeness of samples (making criterion 1 more challenging) (Dennis and Thomas, 2000; Tulloch and Szabo, 2012) and possibly low detectability for many species, because not all surveyors from the wider pool are experts at detecting and identifying species (detrimentally affecting criterion 3) (Bird et al., 2014; Kelling et al., 2015; Johnston et al., in press). Further, they may preferentially record some species over others (thus compromising criterion 4) (Boakes et al., 2010). Citizen science monitoring schemes therefore traditionally have a trade-off between number of participants and ability to provide high-quality data to estimate biodiversity trends. However, the trade-off is not as stark as it was previously and schemes similar in standard to the UK BBS can now be contemplated in many more countries, and on more taxa. One change is that more citizen scientists are now available, because inexperienced wildlife watchers can access information on the web to help identify and record species. Another is that good quality digital photos can be taken with relatively inexpensive and small cameras. Such photos can be submitted to an online forum or app specialising in identification of species in the taxon of interest, and either experts or other users of the forum (an example of 'crowdsourcing') can help with identification. In the latter case, reliability of identifications can be assessed according to number of respondents and the degree of agreement (e.g. ispotnature.org). There is also a rapidly-developing ability for automatic identification of species in photographs (e.g. http://merlin.allaboutbirds.org/photo-id/).

The feasibility of large-scale monitoring schemes is also improving with advancing technology. For example camera-traps are being increasingly used to record terrestrial mammals, and methods are being developed to convert such data to abundance estimates, using spatially-explicit capture-recapture methods (Borchers and Efford, 2008) or distance sampling methods (Howe et al., 2017). Acoustic detectors can be used in a similar way, and have considerable potential for example for surveys of birds or amphibians in difficult-to-survey habitats such as rain forest (Leach et al., 2016) and for nocturnal species such as bats (e.g. Britzke et al., 2011; Walters et al., 2012). In inaccessible terrestrial environments, acoustic detectors could be placed and collected by drones. The acoustic approach will become more feasible as software is developed to pick out relevant noises from the recordings and automate species identification (Walters et al., 2012; Stowell and Plumbley, 2014; Kalan et al., 2015). Again, crowdsourcing might be at least an interim solution to identifying species in large numbers of audio recordings or images (Swanson et al., 2015). These methods often have only a small number of sensors (e.g. camera traps or acoustic detectors) and locations are often non-random. This makes it challenging to meet criteria 1 and 2. However, the passive monitoring devices record for a long period of time and without the presence of humans, so these methods have high detectability for many vocalising (acoustic) or moving (photographic) species, fulfilling criterion 3.

Swanson et al. (2015) show that it is feasible to carry out a camera trap survey over a large area – Serengeti National Park in Tanzania in the case of that study. As the technology advances and costs reduce, it becomes feasible to implement monitoring surveys in countries with limited resources, especially when the surveys are supported by international agencies. To implement a scheme to monitor regional biodiversity trends in a community across a broad spatial extent, a modest number of sensors will often be sufficient. If a pilot survey is conducted, power analyses can be applied to estimate the number of sensors required for a given level of precision (Legg and Nagy, 2006). If there is interest in quantifying how temporal trends vary spatially, or how they vary by habitat, then substantially more sensors are required.

In marine environments, acoustic detectors can be placed on underwater gliders (autonomous underwater vehicles), which require very little power and can travel thousands of kilometres. Alternatively acoustic detectors can be fixed to drifters, which drift through the ocean with the current. Given the difficulty of following designed transects, spatio-temporal modelling will be required to estimate trends over the survey region from data gathered from such platforms, requiring potentially complex model-based solutions to meet criterion 1 for representativeness.

High-resolution photographic imagery is also becoming a new and useful technique for monitoring biological communities. This approach is already widely used with piloted aircraft (e.g. Buckland et al., 2012b). As long-range drones become more widely available, and restrictions on their use relaxed, they might be used to conduct strip transect surveys, recording high-resolution imagery. This method is particularly useful for animals that stand out from their environment, for example bears in the tundra (e.g. Stapleton et al., 2016), seals hauled out on coastlines (e.g. Conn et al., 2014), or mammals and birds in marine environments (e.g. Johnston et al., 2015). Satellite images also have potential for monitoring biodiversity (e.g. Convertino et al., 2012). Software to identify sections of images that have objects of interest, and possibly also to provide species identification (e.g. Mata-Montero and Carranza-Rojas, 2016; Martineau et al., 2017), make it more feasible to process the data from surveys that generate large numbers of high-resolution images. With this surveying method, it is easy to achieve large samples (criterion 2) and detectability can be high for many species if all images are processed and a high-resolution camera is used (criterion 3), but this method is most suitable in open and uniform habitats, for example marine environments; it cannot be used to achieve a representative sample of heterogeneous terrestrial environments (criterion 1) due to large differences in detectability by habitat.

Another technology that opens up new possibilities for biological monitoring is analysis of environmental DNA (eDNA). Small amounts of DNA are naturally released into the environment, for example from scales, skin, saliva or faeces. Modern techniques enable samples from environments to identify the species that have recently been present and therefore create a species list and potentially species abundances for the site. To date this technique has been most useful in freshwater environments (Thomsen et al., 2012), but there is also scope for it to be used in marine and terrestrial habitats (Foote et al., 2012, Bohmann et al., 2014). In suitable environments, this method of sampling has high detectability of many species (fulfilling criterion 3), but for some taxa little is known about the uncertainty in species identification (Somervuo et al., 2017) and therefore the relative numbers of false presences and false absences. eDNA sampling also has potential for estimating abundance through capture-recapture of individual genetic identifiers. eDNA is not yet conducted at large enough scales to achieve a high sample size across a large area (challenging criterion 2); however the development of technology or use of citizen science (Biggs et al., 2015) may make this more feasible in the future. Other issues that would need to be considered are that DNA can be transported over long distances (Deiner and Altermatt, 2014), and may still be detected after many decades (Yoccoz et al., 2012).

There is a final category of biodiversity assessment that, unlike those above, does not require identification of species. Technology is providing methods to assess diversity in acoustic landscapes. This is potentially a powerful technique, for example estimating phylogenetic biodiversity (Gasc et al., 2013) or ecological condition (Tucker et al., 2014) without identifying individual species and only assessing the complexity of the overall acoustic soundscape. These methods potentially yield large sample sizes and it would be possible to create representative sampling strategies (criteria 1 and 2). However the detectability criterion is harder to assess, because often it will be challenging to know which portion of a biological community is being assessed with the acoustic soundscape and it is also difficult to know whether or not this sampled community is representative of the entire community (criterion 4). For example, a family of birds with complex song and mimicry may be over-represented in an analysis of acoustic biodiversity from soundscapes. Until more research is conducted, it is difficult to assess whether this monitoring method will meet criterion 4.

Issues for monitoring biodiversity of large marine regions differ from those for terrestrial regions and in many cases are more complex, due to the issues of low detectability and paucity of marine citizen scientists. The one marine environment in which citizen science programmes have contributed substantially is monitoring coral reefs. For example, programmes have utilised volunteer divers to monitor corals (e.g. Done et al., 2017) and volunteers on computers to classify photographs of reefs (e.g. Parkinson et al., 2016; Raoult et al., 2016). However, in offshore marine environments, survey ships are needed; even if the species of interest can be detected from the air, survey aircraft do not have the range to survey large regions. Given the costs of large-scale surveys, any assessment of biodiversity trends is likely to be an addition to surveys that are being conducted for another purpose. For example, shipboard line transect surveys were conducted over two decades in the Eastern Tropical Pacific, to estimate trends in stocks of dolphins affected by the tuna purse-seine fisheries (Gerrodette et al., 2008), and international trawl surveys have been conducted for over four decades in the North Sea to assess abundance of commercial fish stocks (http://ocean.ices.dk/Project/IBTS/). Both databases offer the potential for estimating biodiversity trends. However, there is a need to develop robust monitoring programmes to assess biodiversity trends of marine fauna (Greenstreet, 2008; Edgar et al., 2016).

4. Estimating biodiversity trends

Biodiversity is a multivariate concept, and any single measure will fail to summarize all the information in the time series of species abundance estimates (Buckland et al., in press). For example, McGill et al. (2015) identify fifteen forms of biodiversity change. While a single headline indicator can be useful for highlighting biodiversity changes for policymakers, analysts tend to compensate for the loss of information when species trends are amalgamated into a composite index by providing additional plots as for example in Fig. 1, which shows estimated trend in biodiversity for priority species in the UK. The left-hand plot shows separate trends for different taxa, allowing more informed interpretation of the headline trend.

Here we focus on the headline measure that is typically used for assessing progress towards CBD targets: the geometric mean of species indices. Usually, the species indices would be a measure of abundance of each species, relative to that species' abundance in a baseline year. The merits of using the geometric mean rather than any of the more classical measures are discussed by Buckland et al. (2011). Because the baseline year typically corresponds to the first year of data, for which sample size is often low, estimation is likely to be improved by smoothing the time series of abundance estimates for each species. This has the added advantage that any zero estimates arising from failing to record a species in a given year are replaced by smoothed non-zero estimates. (A geometric mean cannot be calculated if any estimate is zero, unless an arbitrary value is added to it.) A further advantage of a smoothed estimate of the trend is that spatial variation in temporal biodiversity trends becomes more evident, as most of the fluctuation arising from sampling error is removed (Harrison et al., 2014; Massimino et al., 2015a).

Turnover measures summarize a different aspect of biodiversity and quantify how community composition is changing. Most turnover measures are based on changing ranges of species, but when interest is in large regions, as for CBD targets, it is difficult to establish when a species becomes extinct or colonizes, and such events are typically fairly rare (Buckland et al., in press). When monitoring provides multi-species data from which abundance can be estimated, we can instead base turnover measures on the changing species proportions in the community (Harrison et al., 2016; Yuan et al., 2016). Such measures are more sensitive to changes arising for example from climate change, because gradual shifts of range will be reflected in changing species proportions long before regional extinctions occur (Massimino et al., 2015b; Harrison et al., 2016).

The precision of biodiversity trend estimates is often calculated using bootstrapping methods (Buckland et al., 2005). When the data arise from a designed and randomised survey, it is natural to resample locations in a way that respects the design. For example in the case of the UK BBS, a stratified random sample of 1km squares is selected. In this case, to generate a bootstrap resample, for each geographic stratum, we would select a sample of the surveyed squares in that stratum with replacement, keeping the sample size fixed. Thus if there were 20 surveyed squares in a given stratum, we would select 20 with replacement from that stratum. In any given bootstrap replicate, some squares are selected more than once, while others are not selected at all. We repeat this process for all strata. The bootstrap resample is analysed for each species in the same way as for the real data, and the whole process repeated a large number of times. The variability in estimated trends from the bootstrap resamples is used to estimate confidence limits for trend estimates (Buckland et al., 2005). However, data are often collated from a range of surveys and the resampling cannot follow the same formal structure. For example, the Living Planet Index (LPI, Loh et al., 2005) has no underlying design, and datasets from multiple sources are used. Thus the bootstrap cannot be implemented in the same way. Instead, we might resample species, or, as there are multiple datasets on many species in the LPI, we might resample datasets. An assumption for this bootstrapping technique (as well as for the main indicator) is that the species set is representative of all the species in the community of interest (criterion 4). For any scheme based on a randomized design, resampling should be based on the sampling units that are randomized; only when this is not an option should resampling of species or datasets be considered. When locations are resampled, inference is restricted to those species included in the analysis. By contrast, when species are resampled, inference is on a wider community that the species are assumed to represent. Because different species may show very different trends, the latter approach tends to generate wider confidence intervals (Buckland et al., 2005).

5. Monitoring programme pitfalls

Various issues arise that might compromise biodiversity trend estimates when monitoring programmes are established, when data are gathered, or when trends are estimated. Several of these issues and their implications for conservation management are discussed here.

5.1 Poor estimation in baseline year

Biodiversity monitoring often relies on measuring trends from an initial baseline year. Examples include the Living Planet Index (Loh et al., 2005) and the UK's Wild Bird Indicators (Gregory and van Strien, 2010). Inaccurate estimates in the baseline year will usually lead to inaccurate estimates of the population trend. Fig. 1 shows estimated trends in relative abundance of priority species included in an indicator used to report on progress with international commitments on biodiversity (Burns and Eaton, 2014). The separate trends for four species groups are shown in the left-hand plot. Two of these (moths and butterflies) each show a 40% drop in abundance from the first year they enter the indicator to the second, for reasons that are unclear. A third group (mammals) shows a 40% increase in the first four years that they are included. In the first two years (1993-1995) the trend is determined solely by the dormouse survey (Burns and Eaton, 2014), so that the estimated trend is unrepresentative of the whole group of mammals (criterion 3). It is evident that measuring a trend relative to a baseline year is problematic if there are large annual fluctuations in an index, or if the baseline year is the first year of a time series, when there are possibly comparability issues until a scheme has 'settled down', or if the baseline year has a low sample size and is therefore subject to greater sampling variation (failure to meet criterion 2), or if the baseline year has a small number of species (failure to meet criterion 4).

Sensitivity to choice of baseline year can be reduced by smoothing the index, for example using generalized additive models (Buckland et al., 2005). Also, the first year of the time series need not be the baseline year; choosing a year for which there are more data will tend to reduce bias and increase precision. Fig. 2 illustrates both strategies; smoothed trends have been fitted to the point estimates, and the high variance in the early roost count indices is not reflected in the later years because a baseline year has been selected for which precision is good (Barlow et al., 2015). Another option is to have say a ten-year moving window, so that the baseline year advances by one each year. Dependence on the baseline year can be removed entirely by estimating the second derivative of the smoothed index. If this derivative is significantly greater than zero for a given year, then this is evidence of a reduction in the rate of loss of biodiversity, or an increase in the rate of gain (Buckland et al., 2005). Harrison et al. (2014) exploited this approach to quantify changes in the UK breeding bird communities.

Poor estimation in the baseline year could impact conservation biology by leading to imprecise trends with wide confidence intervals. This could lead to biodiversity declines being overlooked, because they are not identified with confidence. Additionally, many conservation applications ignore the uncertainty around estimates of species trends, and imprecise trends can mislead when they are assumed to be known with certainty. This could lead to false classification of the status of species or communities.

5.2 Species selection

A non-representative set of monitored species can lead to estimates of biodiversity that do not accurately reflect the true community biodiversity. Fig. 1 illustrates the issue of species selection. The 'priority species – relative abundance' indicator features 213 species but is intended to represent 2890 priority species (many of which are priority species due to population declines). The selection of the 213 species is largely determined by availability of time series of estimates. The 2890 species include a wide range of plants, vertebrates and invertebrates, whilst the indicator is dominated by birds and moths (Fig. 3). Fig. 1 shows that different taxa have quite different trends, and the long-term decline in the overall index is largely driven by moths. This can be seen as a failure to meet criterion 4, as the species of interest are not well monitored by the survey methods. This indicator clearly cannot be considered a good guide to trends across the full set of 2890 priority species. To correct for this biased sampling of species, we can theoretically weight the index to reflect the proportion of species from each taxon that are included (Buckland et al., 2012a). However, the index cannot reflect trends within the taxa that are not included in the index (e.g. plants), and also there is no guarantee that within a taxon, those species included in the index are representative of the full set of species from that taxon.

Due to the limitations of the above index, a second index is produced in the UK that is based on occurrence data. The requirement for only occurrence data (rather than abundance data) makes it possible for a wider range of species to be included in the index. The 'priority species – frequency of occurrence' index is a composite indicator of 111 species, including: bees, wasps, ants, dragonflies, grasshoppers and related insects, ground beetles, moths, bryophytes, and freshwater fish. Using occurrence rather than abundance allows a more representative species sample, but the metric now measures a different quantity. This has not stopped authors from taking the geometric mean of trends based on the two different strategies, despite the difficulty in interpreting resulting trend estimates (van Strien et al., 2016). This is an example of the kind of trade-offs that are often made in producing biodiversity indices.

The conservation implications are most severe if the trends of the monitored species are more positive than the trends of the other species and the index of biodiversity will therefore be positively biased. Conservation measures may be designed to target the species included in the indicator. Particularly in situations with limited resources, it may be politically strategic to target efforts towards species in which the impact of conservation policies will be measurable. To improve the representativeness of multi-taxa indices, we recommend that at least a few species are monitored from each taxon. This would enable the index to be weighted to account for biased species selection; this is not possible if there are none or very few species monitored from a given taxon. Simulations could be used to assess how many species of each taxon should be monitored to achieve the desired precision in the weighted index.

5.3 Monitoring known colonies

Monitoring species that occur in large colonies can be challenging. Monitoring sites are usually selected based on known colonies, which can introduce an element of bias into the estimates of population trend if there is turnover of colonies. As existing colonies become extinct and new colonies establish, we see a downward trend in surveyed colonies as some of them are lost, but we fail to measure the corresponding increase resulting from the appearance of new colonies. This can lead to negatively biased estimates of population trends as declining sites or sites that go extinct tend to be over-represented, while increasing sites or newly-established sites are under-represented. This is a failure to meet criterion 1 as the colonies monitored at any given time are not representative of the whole population.

In the case of bat monitoring in the UK, bias of this type can arise for summer roost counts. With the development of bat detectors, many species are now monitored by field survey, and provided representative sites are surveyed, these surveys are free of such bias. There is however the potential for positive bias in such surveys, as technological advances in bat detectors increase detectability. Barlow et al. (2015) included detector type in their models of trend, to adjust for such advances.

'Roost-switching' refers to when some or all bats in a roost move to another location. This can cause bias in the estimated trends (Barlow et al., 2015). We pick out results for the common pipistrelle bat in the UK, which has been surveyed using summer roost counts (affected by roost-switching) and field surveys (which are not affected). The smoothed index for common pipistrelles shows an 82% increase from 1999 to 2015 based on field counts using bat detectors, while similar analyses of summer roost counts show a decline of 58% (Fig. 2); the respective confidence bands indicate that the difference is much greater than can be explained by chance. There are three species with both field survey and roost count data, and the soprano pipistrelle shows a similar discrepancy to the common pipistrelle, while any effect for the serotine is relatively small. Pipistrelle bats have a high degree of roost-switching in the UK, which is particularly likely to lead to non-representative monitoring and a biased trend estimate. For species that have a high degree of fidelity to summer roost sites (such as greater and lesser horseshoe bats), bias introduced from monitoring known roosts will be small.

The potential pitfall of monitoring known colonies is a failure of the monitoring methods to meet criterion 1 – at any one point in time, the trends within monitored sites are not representative of the trends at all sites, because colony abandonments are monitored whilst colony establishments are often missed.

The conservation implications of monitoring known colonies are that estimated population trends may be negatively biased and conservation resources may be focussed on species or regions where they are not needed. This highlights the need for representative sites. Even a small sample of representative sites may be sufficient to assess the degree of bias in a scheme based on monitoring known colonies. It is also important to add new sites when they are first identified, particularly if they are identified early in their growth. However, adding previously unmonitored established sites should be done with caution. Introducing new sites only when they have high abundance is statistically known as 'preferential sampling' (e.g. Shaddick and Zidek, 2014). Further, even if at the outset of a sampling programme, a simple random sample of colonies is selected for monitoring, a strategy of monitoring those colonies for as long as they exist will generate downward bias in trend estimates unless there is a mechanism for adding a representative sample of new colonies each year.

5.4 Measuring trends at atypical locations

There are many examples of surveys in which the locations sampled are unlikely to exhibit trends that are representative of the community for which inferences are required, violating criterion 1 of the criteria for designing monitoring surveys. The Living Planet Index is taken as an indicator of global biodiversity trends. Its geographic coverage is shown in Fig. 4, from which it is evident that some regions are very over-represented relative to others. McRae et al. (2017) developed a diversity-weighted version of the index in an attempt to eliminate taxonomic and geographic bias, 'by accounting for the estimated number of species within biogeographical realms, and the relative diversity of species within them' (Fig. 5). While their analysis is a large step in the right direction, many subjective decisions are made in determining the weighted index, and the large difference in the two trend estimates of Fig. 5, with widely-separated confidence bands, should be seen as a warning – other plausible choices of weighting may generate quite different trend estimates. For example, the Palearctic is a single geographic stratum in their weighted analysis, yet sampling in this region is heavily biased towards the western quarter of the region (i.e. Europe). Within regions, there is more sampling in areas of higher population density, where anthropogenic effects on biodiversity are likely to be greater, yet the weighted analyses assume representative sampling of locations within a region. Similarly, in oceanic strata, sampling is heavily biased towards continental shelves, where the effects of commercial fisheries, disturbance and pollution are likely to be greater than in the open ocean. As a consequence, it seems unlikely that even the biodiversity-weighted trends shown in Fig. 5 accurately quantify loss of biodiversity globally. In principle, model-based methods could be developed to attempt to adjust for these spatial biases.

The survey routes in the North American Breeding Bird Survey follow tracks and roads, a form of preferential sampling. Peterjohn and Sauer (1994) estimated trends of woodland birds from these data. They concluded that, while most woodland communities were doing reasonably well, in the period 1982-1991, Neotropical migrants had fared badly. While this may well be true, we cannot have full confidence in the conclusion because sampling is along roads and tracks where disturbance and loss of habitat are likely to be greater than for more representative locations (Keller and Scallan, 1999). Further, increasing traffic volumes and noise over time may lead to reduced densities along the routes (e.g. Summers et al., 2011), and reduced detectability of singing and calling birds (e.g. Pacifici et al., 2008).

The UK's Butterfly Monitoring Scheme (BMS) is another example in which atypical locations are monitored. Sites tend to be selected because they provide good butterfly habitat, and then transects are placed through the best habitat within the sites. This might bias trends either way. First, trends in abundance may be more favourable in the best sites, which are often protected and managed for conservation, than in the wider countryside. Second, if transects are placed through the best habitat within each site, and the location of the best habitat changes over time while the transects are fixed, then declines may be observed in the counts which are not indicative of trends within the sites. Similarly, if the best sites are selected for monitoring, and there is turnover in which the best sites are, monitored sites might show declines, while comparable, unmonitored sites might show increases (similar to the colony count issue outlined above). These effects are examples of 'regression to the mean': top-ranked sampling units tend to fall in the ranks, while low-ranked units tend to improve on average. If units are selected at random, this does not bias trends, but if top-ranked units are more likely to be sampled, it does. Furthermore, if citizen scientists surveying sites that are no longer good sites are more likely to stop surveying, and new participants are more likely to join the scheme at good sites, this could exacerbate the regression-to-the-mean effect.

In recognition of the possible non-representativeness of BMS sites, the Wider Countryside Butterfly Survey was established in 2009, in which two 1km transects are surveyed in selected 1km squares (Brereton et al., 2011). The number of sites surveyed annually is in the high hundreds, roughly the same as for BMS (Roy et al., 2015). The squares are selected according to a stratified random sampling scheme, and the idealized route is independent of habitat. Roy et al. (2015) compared the two schemes, and found broad agreement in trends, although two species showed significant trends in opposite directions for the two schemes. They found that precision was appreciably higher for BMS, which was attributed to the fact that BMS involves a number of visits each year spread through the whole season.

There are two potential pitfalls of the above approach for wider countryside monitoring. Firstly, most squares were originally selected for the UK's Breeding Bird Survey, and observers were given the option of also recording butterflies (on additional visits). Thus there is an element of selection (observers may be less inclined to record butterflies in sites where few butterflies occur for example), thus compromising the random design. Further, it is usually not possible to follow the idealized route, and the transects are shifted for example so that they run along field edges, rather than through crops. This may generate relatively little bias for bird count trends, given that birds out to 100m either side of the transect are included in analyses, but for butterflies, a 2m-wide box is used, and so butterfly counts are heavily biased towards edge habitats. The degree of bias depends on the habitat; in grazed grassland and in natural or semi-natural habitats, it is often possible to follow the idealized route quite closely, while in arable crops, it is not.

Thus the BMS constructs population trends using data from atypical locations, which does not meet criterion 1 of representative sites in the monitoring scheme. However, Dennis et al. (2013, 2016) have analysed the data from the BMS accounting for phenology of the butterfly flight period and missing visits, using model-based approaches to account for some of the biases in the data and the potential impact of phenological changes on trends.

For circumstances in which trends are similar across the whole area, then the bias in estimated trends from sampling non-representative sites may be small and the conservation implications correspondingly small. Dennis et al. (in press) for example found that trends estimated from the UK Big Butterfly Count data were consistent with those estimated in the BMS, despite the fact that both schemes survey a non-representative set of sites. However, for many taxa, the bias produced by non-representative sites is likely to vary across species and regions and it is difficult to generalize concerning the likely impact. To address the issues inherent in non-representative samples, various options are available. In some cases, smaller representative samples could be used for comparison; however for estimating trends, it is important that these cover the same time frame as the whole sample. Alternatively, model-based approaches could be used to account for bias introduced by non-representative samples (e.g. Stolar and Neilsen, 2015; Kéry et al., 2010; Dennis et al., in press).

5.5 Reliance on relative measures of abundance

When quantifying biodiversity trends of a large region, ideally measures would be based on absolute estimates of abundance in the region. If resources are insufficient to allow reliable estimation of abundance, sample counts for a given amount of effort (e.g. time in the field, number of traps or length of transect) are often assumed to be proportional to abundance. When detectability varies by species, this may generate bias in trends if biodiversity measures based on species proportions are used, but not when the geometric mean of relative abundance is used (Buckland et al., 2010). However, if there is a trend in detectability over time, and it is not modelled, then estimates of trend will be biased. This has been identified as a problem in the North American BBS, for which the average age of observers has increased appreciably since its inception in the early 1960s. Farmer et al. (2014) found substantial evidence of declines in detectability with observer age, concluding that observer aging can negatively bias long-term monitoring data for some species. They recommended that survey designers and modellers should account for observer age. Other possible causes of bias in trends arising from changing detectability include habitat succession, improvements in technology for detection over time (e.g. improved bat detectors, digital images or acoustic recordings), phenological changes (e.g. earlier leaf unfolding), species behavioural changes (which might be linked to phenological changes), and observer learning (Kelling et al., 2015).

If relative abundance trends are assumed to reflect trends in absolute abundance, conservation managers will be misled when detectability changes over time. Ideally, field methods would be used that enable the estimation of detectability, for example distance sampling or occupancy modelling. In these cases, detectability can be incorporated into trends and changes to detectability tested and accounted for in subsequent trends. However, monitoring data often do not allow the estimation of detectability (Watson, 2017). In these cases, the effect of changing detectability can be partially accommodated in the model by including covariates that describe factors associated with detectability. For example, improving acoustic technology could be included by a covariate describing the equipment type (e.g. Barlow et al., 2015), or the aging of observers could be modelled by a covariate of observer age. Such modelling will go some way towards accounting for the effect of changing detectability, even in a model where detectability is not explicitly estimated. The effect of changing technology or a pool of observers that age (Farmer et al., 2014) or learn (Kelling et al., 2015) will be particularly important in surveys that span a long time frame as the changes are likely to be more significant. Often the assumption of constant detectability will be reasonable (e.g. Newson et al., 2013). In summary, when there are known or suspected sources of variation in detectability, the best course of action is to estimate detectability, and failing this, to include variables in the model that describe the key sources of variation.

5.6 Monitoring sample plots within colonies

Here we present another problem with monitoring densely populated colonies. The example of bats above surveyed all individuals within a roost. However, colonies of breeding seabirds are often surveyed by monitoring sample plots within the colony (Walsh et al., 1995). Subject to overcoming the difficulty of counting random plots, and defining plots on what may be very steep and irregular ground, the method should give unbiased estimates of time trends, if population size changes as a result of increasing or decreasing density across the colony. However, there is a potential pitfall if density stays constant and population size changes by expansion or contraction of the colony. In this case, colony expansion will not be detected by sample plots that were designed on the previous colony extent. It is therefore important that the sampling scheme is designed to expand with the colony, maintaining the same sampling rate across the colony (Walsh et al., 1995). If there are sufficient surveyed plots randomly placed across the colony, contraction will not create bias, as the counts in plots that are beyond the new boundaries are simply recorded as zero, thus reflecting the decline. However, if plots are selected that are entirely internal to the old colony (i.e. avoid the colony edge), the failure to sample colony edges means that contraction will take longer to detect. Haines and Pollock (1998) outline a method for surveying eagle nests where a larger area is randomly sampled to detect new nests and assess the completeness of a more focussed survey. For some taxa and species, a similar method could be employed, where an area larger than the colony is defined at the start of the survey and randomly sampled each year, in addition to the regular surveys, in order to detect colony expansions. For example, for species expanding as a result of climate change, a standardised survey such as BBS could be used as the random samples from a larger area.

Colony contraction is easier to accommodate in a good sampling design and is more important for conservation. Plots that used to be within the colony can be monitored as the colony is declining (and recorded as zero count once they are entirely outside the colony). Colony expansion is more difficult to accommodate in hindsight. Plots should be added as soon as possible and ideally in anticipation of colony expansion. In the absence of adequate planning, these situations will usually lead to population increases that are not fully captured by the estimated trend. This error will usually not be a problem for conservation decisions, for which failure to detect declines and falsely detecting declines are more critical errors.

5.7 Over-ambitious objectives

Site-based biodiversity monitoring often focuses on understanding communities, so that large volumes of detailed data are recorded, such as at Barro Colorado Island, which was 'constructed specifically to allow long-term observation of tropical organisms: their complex behaviors, life histories, population dynamics, and changing species composition' (Raby, 2015). This is not achievable when the objective is to monitor regional or national biodiversity trends. First, a large sample of representative sites is required (criterion 1). Second, field methods must be sufficiently simple for large numbers of volunteers to be able and willing to record useful data (criterion 2). Thus the focus should be exclusively on gathering data that allow reliable quantification of species-specific trends in abundance (absolute or relative) within the region. Waldon et al. (2011) called for the adoption of a simple sampling scheme that can be applied throughout a region for monitoring tropical forests. While the details of their proposals are subject to debate (Harrison et al., 2012), the principles are sound. A regional scheme does not need to be capable of reliably quantifying biodiversity trends at each sampled site. Instead, their importance arises from the fact that they are representative sites of the region, and enable regional trends to be accurately quantified.

There are several methods to assess whether a sampling scheme is capable of accurately monitoring the biodiversity within a region. In straightforward scenarios, we advocate the use of power analyses. However, in complex situations it may be challenging to produce a realistic power analysis. Several assumptions are required to use the biodiversity trend as indicative of the trend in a wider region or a wider set of species. We suggest that these assumptions are explicitly stated and that trends are interpreted with caution and with regard for the uncertainty in the trends. This will enable conservation managers who use these trend estimates, to consider the implications of violated assumptions. Overall, to ensure that conservation management is based on honest assessments of biodiversity, we promote candid presentations of the assumptions used to extrapolate trends to large regions or sets of species and explicit presentation of uncertainty.

6. Discussion

In some cases, regional monitoring involves large-scale surveys, such as the ship surveys conducted in the eastern tropical Pacific (Gerrodette et al., 2008). More commonly, regional monitoring is achieved by conducting small-scale surveys at a number of locations spread through the region, as in the UK Breeding Bird Survey. Technology that to date has mostly been used for small-scale surveys, such as camera traps, acoustic detectors and drones, is improving and becoming more accessible, and will be key to generating more reliable data from large-scale regional surveys. Technology also provides practical options for conducting citizen-science regional surveys on a range of taxa.

Rigorously-designed monitoring schemes will usually produce estimated trends in biodiversity that have low bias and good precision. However, often monitoring schemes are compromised in their sampling design; whilst it is still possible to generate trend estimates, their interpretation is much more challenging, and the implications for biodiversity often unclear. A scheme that fails to meet only one of these criteria might, depending on the objectives and the nature of the failure, have unusable biodiversity trends with extreme bias. Alternatively, a scheme may fail several criteria, and yet still be useful with respect to its stated objectives. As the impact of failing each criterion will vary considerably with objectives and situations, it is important to assess each survey design and investigate the violation of assumptions and power of the design for the purposes of the target biodiversity monitoring.

There is inevitably a trade-off between ideal sampling designs and designs that are realistic and achievable. At one extreme is the ideal of a large scheme, with many sampling locations selected according to a randomised design, and with adequate resources and expertise to ensure that sound data are collected on all species, or a representative sample of species, in the community of interest. This ideal must be assessed against reality. Which compromises are likely to have small impacts and retain the fundamental principles of sampling, and which compromises would result in a scheme that simply is not fit for purpose? The answer to this will vary according to circumstances. It may be that a design-based inference scheme is too unreliable given the compromises (such as sampling non-random sites, or a non-representative set of species) that must be made. In this case, can model-based methods be implemented to eliminate, or at least reduce, the bias present in design-based trends? For example, if detectability is affected by the amount of effort put into sampling, and it is not possible to ensure that the same sampling effort is carried out at each site, modelling detectability as a function of effort should reduce bias. Occupancy modelling methods have been used to good effect to estimate distribution trends from opportunistic citizen science data (Kéry et al., 2010; van Strien et al., 2013). Walker and Taylor (2017) used binomial generalized linear mixed-effects models to estimate trends in bird numbers from the North American citizen-science bird observation network, eBird (Sullivan et al., 2009). When proposing new surveys, we advocate the use of simulations, power analyses and advice from statisticians experienced in survey design. Together with very clear survey goals, these mechanisms will assist in assessing whether a proposed monitoring programme meets the criteria outlined above and whether it will be fit for purpose.

Biodiversity loss is often considered to be more rapid in developing countries (although Rodrigues et al. (2014) identify substantial loss in two of the most developed nations: the United States and Australia). The best schemes for monitoring biodiversity are mostly in developed countries that have adequate resources devoted to monitoring. Proponents of improved schemes are frequently criticised for failing to recognise the realities faced in developing countries, or the difficulties of monitoring more challenging taxa, perhaps with access to very few experts. As noted by Yoccoz et al. (2003), in countries with fewer financial resources, it is more critical that monitoring schemes are efficiently designed for the target objectives. Further, many opportunities are now opening up through technological developments. Consider for example a small group of enthusiasts who wish to set up a recording system for butterflies in a developing country. It is now a simple matter to set up a website that large numbers can access. Cameras are now ubiquitous in mobile phones, and it is possible for volunteer contributors to take adequate photos, which can be uploaded to the site. The site can provide reference material and photo galleries, and the group of enthusiasts can tutor contributors in identification. If the volume of submissions becomes too large, then the wider community can be called upon to identify species in images. If sufficient interest is generated, then participants can contribute to more formal surveys, for example walking transects or setting baits, with data entered online.

If such an approach is not feasible, technology might offer alternatives. For example acoustic detectors could be deployed across a region, possibly using drones where access might otherwise be problematic. If visual images are likely to give better data, then camera traps might be deployed instead. Automated identification of individuals from images or recordings would substantially reduce the cost of processing the data, or interested individuals from around the world might be trained online, and allocated images or recordings to process. As the statistical models adapt to new technologies, and become more sophisticated, then reliable inference on biodiversity trends will increasingly become feasible even for many of the more difficult taxa in remote parts of the world.

The world is currently in the middle of a biodiversity crisis, with substantial reductions in biodiversity in many regions (Butchart et al., 2010). To understand the changes in biodiversity and develop conservation programmes that will be suitable to mitigate or reverse the losses, it is critical to have good quality surveys that produce reliable trends in biodiversity. Although the number of monitoring programmes across the world is increasing rapidly, many of these do not produce trends that are robust or representative. Survey design can often be overlooked or rushed, yet we have demonstrated here that good survey design is critical to producing robust biodiversity indicators. Poorly designed surveys can result in indices that are substantially different from the true underlying trends. The five key criteria presented here are guidelines for those designing new surveys. We also present suggestions for analysing data from sub-optimal surveys, which are the only data available in many regions of the world and for many species groups. Robust indicators of biodiversity can only be produced from good surveys and appropriate and careful analysis.

Acknowledgements

This work is based partly on Defra-funded research project BE0102 and available on the Defra Science website. We thank Katherine Boughey of the Bat Conservation Trust and four reviewers for comments that led to a much-improved paper.

References

Barlow, K.E., Briggs, P.A., Haysom, K.A., Hutson, A.M., Lechiara, N.L., Racey, P.A., Walsh, A.L. and Langton, S.D. 2015. Citizen science reveals trends in bat populations: The National Bat Monitoring Programme in Great Britain. *Biological Conservation* **182**, 14-26.

Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R.A., Foster, J., Wilkinson, J.W., Arnell, A., Brotherton, P., Williams, P. and Dunn, F. 2015. Using eDNA to develop a national citizen science-based monitoring programme for the great crested newt (*Triturus cristatus*). *Biological Conservation* **183**, 19-28.

Binder, D.A. and Hidiroglou, M.A. 1988. Sampling in time. Pp. 187-211 in Handbook of Statistics, Vol 6. Elsevier Science Publishers, North Holland.

Bird, T.J., Bates, A.E., Lefcheck, J.S., Hill, N.A., Thomson, R.J., Edgar, G.J., Stuart-Smith, R.D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J.F., Pecl, G.T., Barrett, N. and Stewart, F. 2014. Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation* **173**, 144-154.

Boakes, E.H., MGowan, P.J.K., Fuller, R.A., Chang-qing, D., Clark, N.E., O'Connor, K. and Mace, G.M. 2010. Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biology* 8(6): e1000385.

Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W. and de Bruyn, M. 2014. Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution* **29**, 358-367.

Borchers, D.L. and Efford, M.G. 2008. Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics* **64**, 377-385.

Brereton, T.M., Cruickshanks, K.L., Risely, K., Noble, D.G. and Roy, D.B. 2011. Developing and launching a wider countryside butterfly survey across the United Kingdom. *Journal of Insect Conservation* **15**, 279-290.

Britzke, E.R., Duchamp, J.E., Murray, K.L., Swihart, R.K. and Robbins, L.W. 2011. Acoustic identification of bats in the eastern United States: A comparison of parametric and nonparametric methods. *Journal of Wildlife Management* **75**(3), 660-667.

Buckland, S.T., Baillie, S.R., Dick, J. McP., Elston, D.A., Magurran, A.E., Scott, E.M., Smith, R.I., Somerfield, P.J., Studeny, A.C. and Watt, A. 2012a. How should regional biodiversity be monitored? *Environmental and Ecological Statistics* **19**, 601-626.

Buckland, S.T., Burt, M.L., Rexstad, E.A., Mellor, M., Williams, A.E. and Woodward, R. 2012b. Aerial surveys of seabirds: the advent of digital methods. *Journal of Applied Ecology* **49**, 960-967.

Buckland, S.T., Magurran, A.E., Green, R.E. and Fewster, R.M. 2005. Monitoring change in biodiversity through composite indices. *Phil. Trans. R. Soc. Lond.* B **360**, 243-254.

Buckland, S.T., Rexstad, E.A., Marques, T.A. and Oedekoven, C.S. 2015. *Distance Sampling: Methods and Applications*. Springer, New York.

Buckland, S.T., Studeny, A.C., Magurran, A.E., Illian, J.B. and Newson, S.E. 2011. The geometric mean of relative abundance indices: a biodiversity measure with a difference. *Ecosphere* **2**(9), Article 100. doi:10.1890/ES11-00186.1.

Buckland, S.T., Studeny, A.C., Magurran, A.E. and Newson, S.E. 2010. Biodiversity monitoring: the relevance of detectability. Pp. 25-36 in *Biological Diversity: Frontiers in Measurement and Assessment.* A.E. Magurran and B.J. McGill (eds). Oxford University Press, Oxford.

Buckland, S.T., Yuan, Y. and Marcon, E. in press. Measuring temporal trends in biodiversity. *Advances in Statistical Analysis – Special Issue on Ecological Statistics*.

Burns, F. and Eaton, M. 2014. UK Biodiversity Indicators, Indicator C4a, Status of UK Priority Species – Relative Abundance. Technical Background Report prepared for the Joint Nature Conservation Committee and Defra. www.jncc.defra.gov.uk

Butchart, S.H., Walpole, M., Collen, B., van Strien, A., Scharlemann, J.P., Almond, R.E., Baillie, J.E., Bomhard, B., Brown, C., Bruno, J., Carpenter, K.E., Carr, G.M., Chanson, J., Chenery, A.M., Csirke, J., Davidson, N.C., Dentener, F., Foster, M., Galli, A., Galloway, J.N., Genovesi, P., Gregory, R.D., Hockings, M., Kapos, V., Lamarque, J.F., Leverington, F., Loh, J., McGeoch, M.A., McRae, L., Minasyan, A., Hernández Morcillo, M., Oldfield, T.E., Pauly, D., Quader, S., Revenga, C., Sauer, J.R., Skolnik, B., Spear, D., Stanwell-Smith, D., Stuart, S.N., Symes, A., Tierney, M., Tyrrell, T.D., Vié, J.C. and Watson, R. 2010. Global biodiversity: indicators of recent declines. *Science* **328**, 1164–1168.

Carlson, M. and Schmiegelow, F. 2002. Cost-effective sampling design applied to large-scale monitoring of boreal birds. *Conservation Ecology* **6**(2), 11. http://www.consecol.org/vol6/iss2/art11

CBD. 2011. Strategic plan for biodiversity 2011–2020 and the Aichi Targets. (http://www.cbd.int/doc/strategic-plan/2011-2020/Aichi-Targets-EN.pdf)

Conn, P.B., ver Hoef, J.M., McClintock, B.T., Moreland, E.E., London, J.M., Cameron, M.F., Dahle, S.P. and Boveng, P.L. 2014. Estimating multispecies abundance using automated detection systems: ice-associated seals in the Bering Sea. *Methods in Ecology and Evolution* **5**(12), 1280-1293.

Convertino, M., Mangoubi, R.S., Linkov, I., Lowry, N.C. and Desai, M. 2012. Inferring species richness and turnover by statistical multiresolution texture analysis of satellite imagery. PLoS ONE **7**(10), e46616.

Deiner, K. and Altermatt, F. 2014. Transport distance of invertebrate environmental DNA in a natural river. PLoS ONE **9**(2), e88786.

Dennis, E.B., Freeman, S.N., Brereton, T.M. and Roy, D.B. 2013. Indexing butterfly abundance whilst accounting for missing counts and variability in seasonal pattern. *Methods in Ecology and Evolution* **4**(7), 637-645.

Dennis, E.B., Morgan, B.J.T., Brereton, T.M., Roy, D.B. and Fox, R. In press. Using citizen science butterfly counts to predict species population trends. *Conservation Biology*.

Dennis, E.B., Morgan, B.J.T., Freeman, S.N., Brereton, T.M. and Roy, D.B. 2016. A generalized abundance index for seasonal invertebrates. *Biometrics* **72**(4), 1305-1314.

Dennis, R.L.H. and Thomas, C.D. 2000. Bias in butterfly distribution maps: the influence of hot spots and recorder's home range. *Journal of Insect Conservation* **4**, 73-77.

Done, T., Roelfsema, C., Harvey, A., Schuller, L., Hill, J., Schläppy, M.L., Lea, A., Bauer-Civiello, A. and Loder, J. 2015. Reliability and utility of citizen science reef monitoring data collected by Reef Check Australia, 2002–2015. *Marine Pollution Bulletin* **117**, 148–155.

Duncan, G.J. and Kalton, G. 1987. Issues of design and analysis of surveys. *International Statistical Review* **55**, 97-117.

Edgar, G.J., Bates, A.E., Bird, T.J., Jones, A.H., Kininmonth, S., Stuart-Smith, R.D. and Webb, T.J. 2016. New approaches to marine conservation through the scaling up of ecological data. *Annual Review of Marine Science* **8**, 435-461.

Farmer, R.G., Leonard, M.L., Mills Flemming, J.E. and Anderson, S.C. 2014. Observer aging and long-term avian survey data quality. *Ecology and Evolution* **4**, 2563-2576.

Field, S.A., Tyre, A.J. and Possingham, H.P. 2005. Optimizing allocation of monitoring effort under economic and observational constraints. *Journal of Wildlife Management* **69**, 473-482.

Foote, A.D., Thomsen, P.F., Sveegaard, S., Wahlberg, M., Kielgast, J., Kyhn, L.A., Salling, A.B., Galatius, A., Orlando, L. and Gilbert, M.T.P. 2012. Investigating the potential use of environmental DNA (eDNA) for genetic monitoring of marine mammals. *PLoS ONE* **7**, e41781.

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**, 61–72.

Gasc, A., Sueur, J., Jiguet, F., Devictor, V., Grandcolas, P., Burrow, C., Depraetere, M. and Pavoine, S. 2013. Assessing biodiversity with sound: Do acoustic diversity indices reflect phylogenetic and functional diversities of bird communities? *Ecological Indicators* **25**, 279-287.

Gerrodette, T., Watters, G., Perryman, W. and Balance, L. 2008. Estimates of 2006 dolphin abundance in the eastern tropical Pacific, with revised estimates from 1986-2003. *NOAA Technical Memorandum NMFS-SWFSC* **422**, 1-39.

Greenstreet, S.P.R. 2008. Biodiversity of North Sea fish: why do politicians care but marine scientists appear oblivious to this issue? *ICES Journal of Marine Science* **65**, 1515-1519.

Gregory, R.D., van Strien, A., Vorisek, P., Gmelig Meyling, A.W., Noble, D.G., Foppen, R.P.B. and Gibbons, D.W. 2005. Developing indicators for European birds. *Philosophical Transactions of the Royal Society of London B* **360**, 269-288.

Gregory, R.D., and van Strien. A. 2010. Wild bird indicators: using composite population trends of birds as measures of environmental health. *Ornithological Science* **9**, 3-22.

Haines, D.E. and Pollock, K.H. 1998. Estimating the number of active and successful bald eagle nests: an application of the dual frame method. *Environmental and Ecological Statistics* **5**, 245-256.

Harrison, M.E., Boonman, A., Cheyne, S.M., Husson, S.J., Marchant, N.C. and Struebig, M.J. 2012. Biodiversity monitoring protocols for REDD+: can a one-size-fits-all approach really work? *Tropical Conservation Science* **5**, 1-11.

Harrison, P.J., Buckland, S.T., Yuan, Y., Elston, D.A., Brewer, M.J., Johnston, A. and Pearce-Higgins, J.W. 2014. Assessing trends in biodiversity over space and time using the example of British breeding birds. *Journal of Applied Ecology* **51**, 1650-1660.

Harrison, P.J., Yuan, Y., Buckland, S.T., Oedekoven, C.S., Elston, D.A., Brewer, M.J., Johnston, A. and Pearce-Higgins, J.W. 2016. Quantifying turnover in biodiversity of British breeding birds. *Journal of Applied Ecology* **53**, 469-478.

Hirzel, A. and Guisan, A. 2002. Which is the optimal sampling strategy for habitat suitability modelling? *Ecological Modelling* **157**, 331-341.

Howe, E.J., Buckland, S.T., Després-Einspenner, M.-L. and Kühl, H.S. 2017. Distance sampling with camera traps. *Methods in Ecology and Evolution* doi:10.1111/2041-210X.12790

Johnston, A., Thaxter, C.B., Austin, G.E., Cook, A.S.C.P., Humphreys, E.M., Still, D.A., Mackay, A., Irvine, R., Webb, A. and Burton, N.H.K. 2015. Modelling the abundance and distribution of marine birds accounting for uncertain species identification. *Journal of Applied Ecology* **52**, 150-160.

Johnston, A., Fink, D., Hochachka, W.M. and Kelling, S.T. In press. Estimates of observer expertise improve species distributions from citizen science data. *Methods in Ecology and Evolution.*

Kalan, A.K., Mundry, R., Wagner, O.J.J., Heinicke, S., Boesch, C. and Kühl, H.S. 2015. Towards the automated detection and occupancy estimation of primates using passive acoustic monitoring. *Ecological Indicators* **54**, 217-226.

Keller, C.M.E. and Scallan, J.T. 1999. Potential roadside biases due to habitat changes along breeding bird survey routes. *The Condor* **101**, 50-57.

Kelling, S.T., Johnston, A., Hochachka, W.M., Iliff, M., Fink, D., Gerbracht, J., Lagoze, C., La Sorte, F.A., Moore, T., Wiggins, A., Wong, W.-K., Wood, C. and Yu, J. 2015. Can observation skills of citizen scientists be estimated with species accumulation curves? *PloS ONE* **10**, e0139600.

Kéry, M., Royle, A.J., Schmid, H., Schaub, M., Volet, B., Häfliger, G. and Zbinden, N. 2010. Site-occupancy distribution modeling to correct population-trend estimates derived from opportunistic observations. *Conservation Biology* **24**, 1388-1397.

Leach, E.C., Burwell, C.J., Ashton, L.A., Jones, D.N. and Kitching, R.L. 2016. Comparison of point counts and automated acoustic monitoring: detecting birds in a rainforest biodiversity survey. *EMU* **116**, 305-309.

Loh, J., Green, R.E., Ricketts, T., Lamoreux, J., Jenkins, M., Kapos, V. and Randers, J. 2005. The Living Planet Index: using species population time series to track trends in biodiversity. *Philosophical Transactions of the Royal Society of London B* **360**, 289-295.

Legg C.J. and Nagy, L. 2006. Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management* **78**(2), 194-199.

Mace, G.M. and Baillie, J.E.M. 2007. The 2010 biodiversity indicators: challenges for science and policy. *Conservation Biology* **21**, 1406–1413.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. and Hines, J.E. 2006. *Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence*. Elsevier.

Maggini, R., Guisan, A. and Cherix, D. 2002. A stratified approach for modeling the distribution of a threatened ant species in the Swiss national park. *Biodiversity and Conservation* **11**, 2117-2141.

Magurran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.M., Elston, D.A., Scott, E.M., Smith, R.I., Somerfield, P.J. and Watt, A. 2010. Long-term data sets in biodiversity research and monitoring:

assessing change in ecological communities through time. *Trends in Ecology and Evolution* **25**, 574–582.

Manly, B.F.J. and Navarro Alberto, J.A. 2014. *Introduction to Ecological Sampling*. Chapman and Hall/CRC

Martineau, M., Conte, D., Raveaux, R., Arnault, I., Munier, D. and Venturini, G. 2017. A survey on image-based insect classification. *Pattern Recognition* **65**, 273-284.

Massimino, D., Johnston, A., Noble, D.G. and Pearce-Higgins, J.W. 2015a. Multi-species spatially-explicit indicators reveal spatially structured trends in bird communities. *Ecological Indicators* **58**, 277-285.

Massimino, D., Johnston, A. and Pearce-Higgins, J.W. 2015b. The geographical range of British birds expands during 15 years of warming. *Bird Study* **62**, 523-534.

Mata-Montero, E. and Carranza-Rojas, J. 2016. Automated plant species identification: Challenges and opportunities. Pp. 26-36 in *ICT for Promoting Human Development and Protecting the Environment*. WITFOR 2016. F. Mata and A. Pont (eds). IFIP Advances in Information and Communication Technology, vol 481. Springer, Cham.

McDonald, T.L. 2003. Review of environmental monitoring methods: survey designs. *Environmental Monitoring and Assessment* **85**, 277-292.

McGill, B.J., Dornelas, M., Gotelli, N.J. and Magurran, A.E. 2015. Fifteen forms of biodiversity trend in the Anthropocene. *Trends in Ecology and Evolution* **30**, 104-113.

McRae, L., Deinet, S. and Freeman, R. 2017. The diversity-weighted Living Planet Index: controlling for taxonomic bias in a global biodiversity indicator. *PLoS ONE* **12**, e0169156.

Newson, S.E., Woodburn, R.J.W, Noble, D.G., Baillie, S.R. and Gregory, R.D. 2005. Evaluating the Breeding Bird Survey for producing national population size and density estimates. *Bird Study* **52**, 42–54.

Newson, S.E., Massimino, D., Johnston, A., Baillie, S.R. and Pearce-Higgins, J.W. 2013. Should we account for detectability in population trends? *Bird Study* 60(3): 384-390.

Nichols, J.D., Hines, J.E., Sauer, J.R., Fallon, F.W., Fallon, J.E. and Heglund, P.J. 2000. A double-observer approach for estimating detection probability and abundance from point counts. *The Auk* **117**, 393-408.

Nielsen, S.E., Haughland, D.L., Bayne, E. and Schieck, J. 2009. Capacity of large-scale, long-term biodiversity monitoring programmes to detect trends in species prevalence. *Biodiversity and Conservation* **18**, 2961-2978.

Pacifici, K., Simons, T.R. and Pollock, K.H. 2008. Effects of vegetation and background noise on the detection process in auditory avian point-count surveys. *The Auk* **125**, 600-607.

Parkinson, J.E., Yang, S., Kawamura, I., Byron, G., Todd, P.A. and Reimer, J.D. 2016. A citizen science approach to monitoring bleaching in the zoantharian *Palythoa tuberculosa*. *PeerJ* **4**, e1815. https://doi.org/10.7717/peerj.1815

Pereira, H.M. and Cooper, H.D. 2006. Towards the global monitoring of biodiversity change. *Trends in Ecology and Evolution* **21**, 123–129.

Peterjohn, B.G. and Sauer, J.R. 1994. Population trends of woodland birds from the North American Breeding Bird Survey. *Wildlife Society Bulletin* **22**, 155-164.

Raby, M. 2015. Ark and archive: making a place for long-term research on Barro Colorado Island, Panama. *ISIS* **106**, 798-824.

Raoult, V., David, P.A., Dupont, S.F., Mathewson, C.P., O'Neill, S.J., Powell, N.N. and Williamson, J.E. 2016. GoPros™ as an underwater photogrammetry tool for citizen science. *PeerJ* **4**, e1960. https://doi.org/10.7717/peerj.1960

Rodrigues, A.S.L., Brooks, T.M., Butchart, S.H.M., Chanson, J., Cox, N., Hoffmann, M. and Stuart, S.N. 2014. Spatially explicit trends in the global conservation status of vertebrates. *PLoS ONE* **9**, e113934.

Roy, D.B., Ploquin, E.F., Randle, Z., Risely, K., Botham, M.S., Middlebrook, I., Noble, D., Cruickshanks, K., Freeman, S.N. and Brereton, T.M. 2015. Comparison of trends in butterfly populations between monitoring schemes. *Journal of Insect Conservation* **19**, 313-324.

Sanderlin, J.S., Block, W.M. and Ganey, J.L. 2014. Optimizing study design for multi-species avian monitoring programmes. *Journal of Applied Ecology* **51**, 860-870.

Shaddick, G. and Zidek, J.V. 2014. A case study in preferential sampling: Long term monitoring of air pollution in the UK. *Spatial Statistics* **9**, 51-65.

Somervuo, P., Yu, D.W., Xu, C.C.Y., Ji, Y., Hultman, J., Wirta, H. and Ovaskainen, O. 2017. Quantifying uncertainty of taxonomic placement in DNA barcoding and metabarcoding. *Methods in Ecology and Evolution*. **8**, 398-407.

Stapleton, S., Peacock, E. and Garshelis, D. 2016. Aerial surveys suggest long-term stability in the seasonally ice-free Foxe Basin (Nunavut) polar bear population. *Marine Mammal Science* **32**(1), 181-201.

Stolar, J. and Nielsen, S.E. 2015. Accounting for spatially biased sampling effort in presence-only species distribution modelling. *Biodiversity Research* **21**, 595-608.

Stowell, D. and Plumbley, M.D. 2014. Automatic large-scale classification of bird sounds is strongly improved by unsupervised feature learning. *PeerJ* **2**, e488.

Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D. and Kelling, S. 2009. eBird: a citizen-based bird observation network in the biological sciences. *Biological Conservation* **142**, 2282-2292.

Summers, P.D., Cunnington, G.M. and Fahrig, L. 2011. Are the negative effects of roads on breeding birds caused by traffic noise? *Journal of Applied Ecology* **48**, 1527-1534.

Sutherland, W.J. 1996. *Ecological Census Techniques: A Handbook*. Cambridge University Press.

Swanson, A., Kosmala, M. Lintott, C., Simpson, R., Smith, A. and Packer, C. 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data* **2**, 150026.

Taylor, B.L., Martinez, M., Gerrodette, T., Barlow, J. and Horvat, Y.N. 2006. Lessons from monitoring trends in abundance of marine mammals. *Marine Mammal Science* **23**, 157-175.

Teilmann, J., Rigét, F. and Harkonen, T. 2010. Optimizing survey design for Scandinavian harbour seals: population trend as an ecological quality element. *ICES Journal of Marine Science* **67**, 952–958.

Thomsen, P.F., Kielgast, J., Iversen, L.L., Wiuf, C., Rasmussen, M., Gilbert, M.T.P., Orlando, L. and Willerslev, E. 2012. Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular Ecology* **21**, 2565-2573.

Tucker, D., Gage, S.H., Williamson, I. and Fuller, S. 2014. Linking ecological condition and the soundscape in fragmented Australian forests. *Landscape Ecology* **29**, 745-758.

Tulloch, A.I.T. and Szabo, J.K. 2012. A behavioural ecology approach to understand volunteer surveying for citizen science datasets. *Emu* **112**, 313-325.

Underwood, F.M. 2012. A framework for adapting survey design through time for wildlife population assessment. *Environmental and Ecological Statistics* **19**, 413-436.

van Strien, A.J., Gmelig Meyling, A.W., Herder, J.E., Hollander, H., Kalkman, V.J., Poot, M.J.M., Turnhout, S., van der Hoorn, B., van Strien-van Liempt W.T.F.H., van Swaay, C.A.M., van Turnhout, C.A.M., Verweij, R.J.T. and Oerlemans, N.J. 2016. Modest recovery of biodiversity in a western European country: The Living Planet Index for the Netherlands. *Biological Conservation* **200**, 44-50.

van Strien, A.J., van Swaay, C.A.M. and Termaat, T. 2013. Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology* **50**, 1450-1458.

Van Swaay, C.A.M., Nowicki, P., Settele, J. and van Strien, A.J. 2008. Butterfly monitoring in Europe: methods, applications and perspectives. *Biodiversity Conservation* **17**, 3455-3469.

Waldon, J., Miller, B.W. and Miller, C.M. 2011. A model biodiversity monitoring protocol for REDD projects. *Tropical Conservation Science* **4**, 254-260.

Walker, J. and Taylor, P.D. 2017. Using eBird data to model population change of migratory bird species. *Avian Conservation and Ecology* **12**, article 4.

Walsh, P.M., Halley, D.J., Harris, M.P., del Nevo, A., Sim, I.M.W. and Tasker, M.L. 1995. *Seabird monitoring handbook for Britain and Ireland*. JNCC, RSPB, ITE, Seabird Group, Peterborough.

Walters, C.L., Freeman, R., Collen, A., Dietz, C., Fenton, M.B., Jones, G., Obrist, M.K., Puechmaille, S.J., Siemers, B.M., Sattler, T., Parsons, S. and Jones, K.E. 2012. A continental-scale tools for acoustic identification of European bats. *Journal of Applied Ecology* **49**(5), 1064-1074.

Watson, D.M. 2017. Sampling effort determination in bird surveys: do current norms meet best-practice recommendations? *Wildlife Research* **44**, 183-193.

Yoccoz, N.G., Bråthen, K.A., Gielly, L., Haile, J., Edwards, M.E., Goslar, T., von Stedingk, H., Brysting, A.K., Coissac, E., Pompanon, F., Sønstebo, J.H., Miquel, C., Valentini, A., de Bello, F., Chave, J., Thuiller, W., Wincker, P., Cruaud, C., Gavory, F., Rasmussen, M., Gilbert, M.T.P., Orlando, L., Brochmann, C., Willerslev, E. and Taberlet, P. 2012. DNA from soil mirrors plant taxonomic and growth form diversity. *Molecular Ecology* **21**, 3647-3655.

Yoccoz, N.J. Nichols, J.D. and Boulinier, T. 2001. Monitoring of biological diversity in space and time. *Trends in Ecology and Evolution* **16**(8), 446-453.

Yoccoz, N.G., Nichols, J.D. and Boulinier, T. 2003. Monitoring of biological diversity - a response to Danielsen et al. *Oryx* **37**, 410.

Yuan, Y., Buckland, S.T., Harrison, P.J., Foss, S. and Johnston, A. 2016. Using species proportion to quantify turnover in biodiversity. *Journal of Agricultural, Biological and Environmental Statistics* **21**, 363-381.

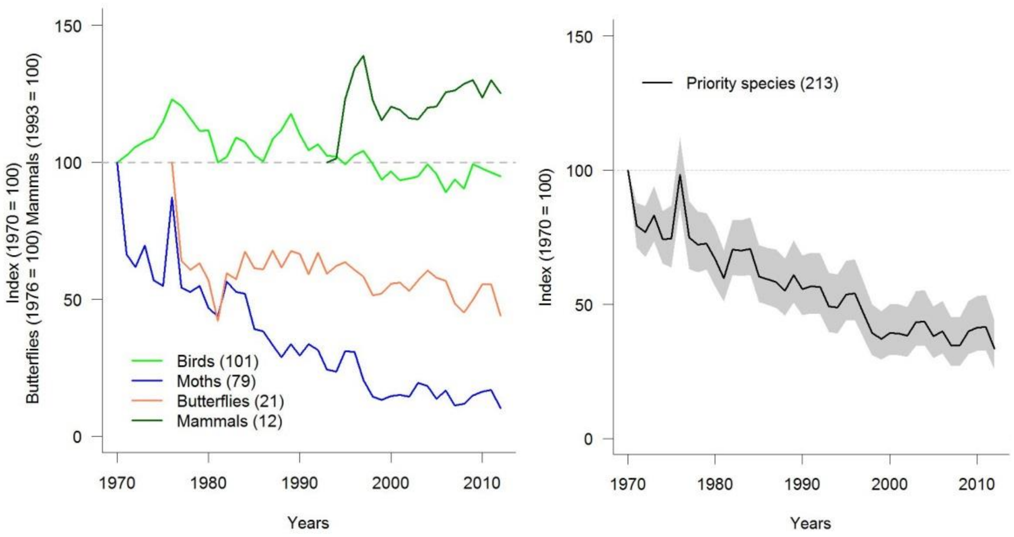

Fig. 1. Index of trends in priority species, split by taxa (left). The 213 separate species trends are combined using a geometric mean of the relative abundance estimates, to form the "priority species – relative abundance" trend used by Defra as a biodiversity indicator (right). Source: Burns and Eaton, 2014.

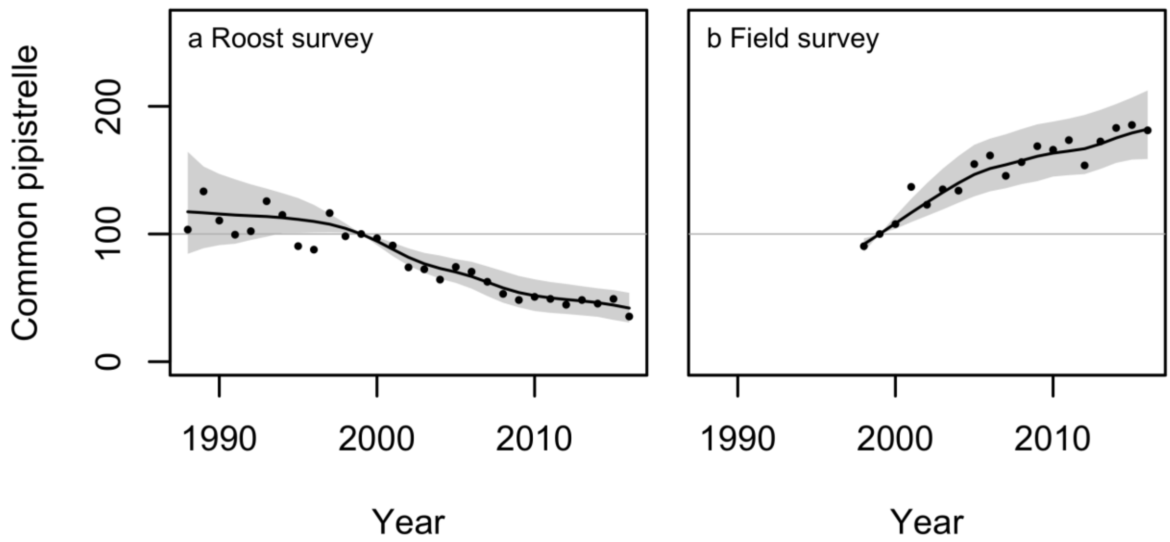

Fig. 2. Method of surveying affects estimated bat population trends of the common pipistrelle in the UK. Smoothed trends have been fitted to the point estimates. The left-hand plot is the trend estimated from roost count data, while the right-hand plot is estimated from summer field survey data. Although the time series of estimates from roost counts starts in 1988, precision was poor in the early years. By taking the baseline year to be 1999, this poor precision does not adversely affect the width of the confidence intervals in later years. This contrasts with the confidence intervals for the field surveys, where the baseline year (again taken to be 1999) is near the start of the time series, and precision is poor on comparisons between that year and subsequent years, resulting in relatively wide confidence intervals. Source: http://www.bats.org.uk/pages/-common_pipistrelle-821.html

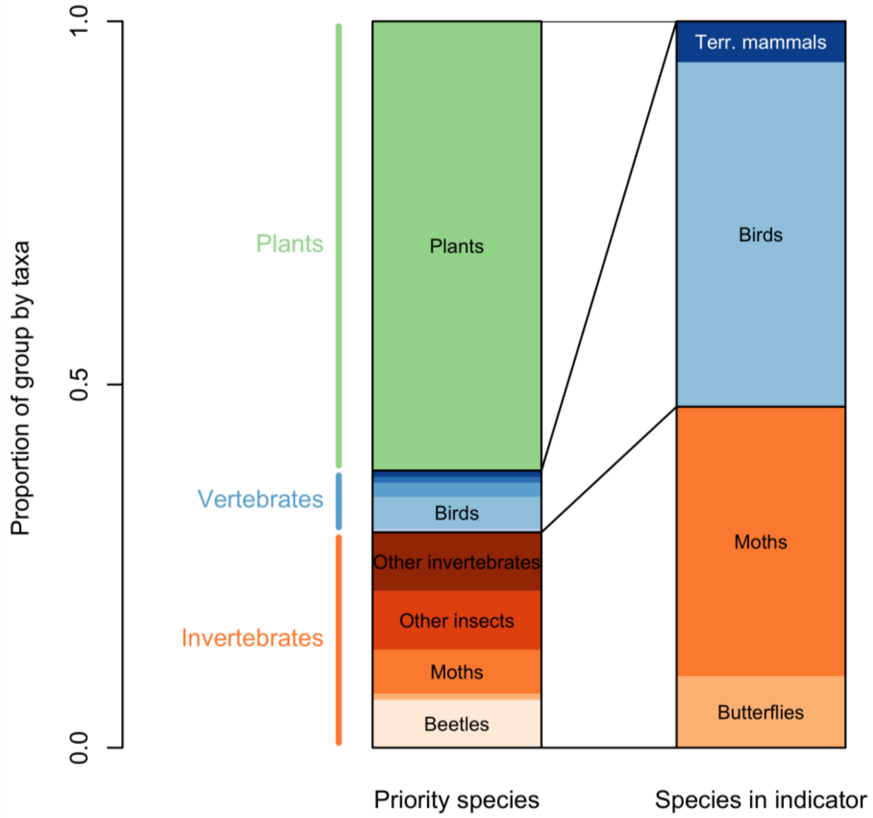

Fig. 3. Proportion of species by taxon in the priority species community, and in the sample used for the relative abundance index.

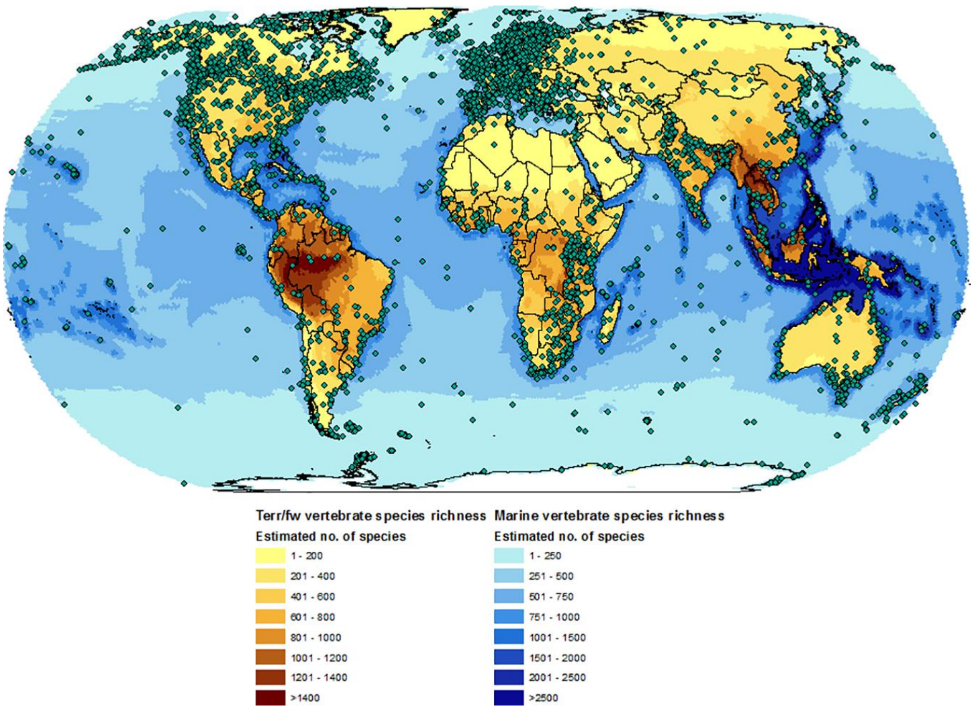

Fig. 4. Global vertebrate richness map overlaid with populations recorded in the Living Planet Database. Reproduced from McRae et al. (2017).

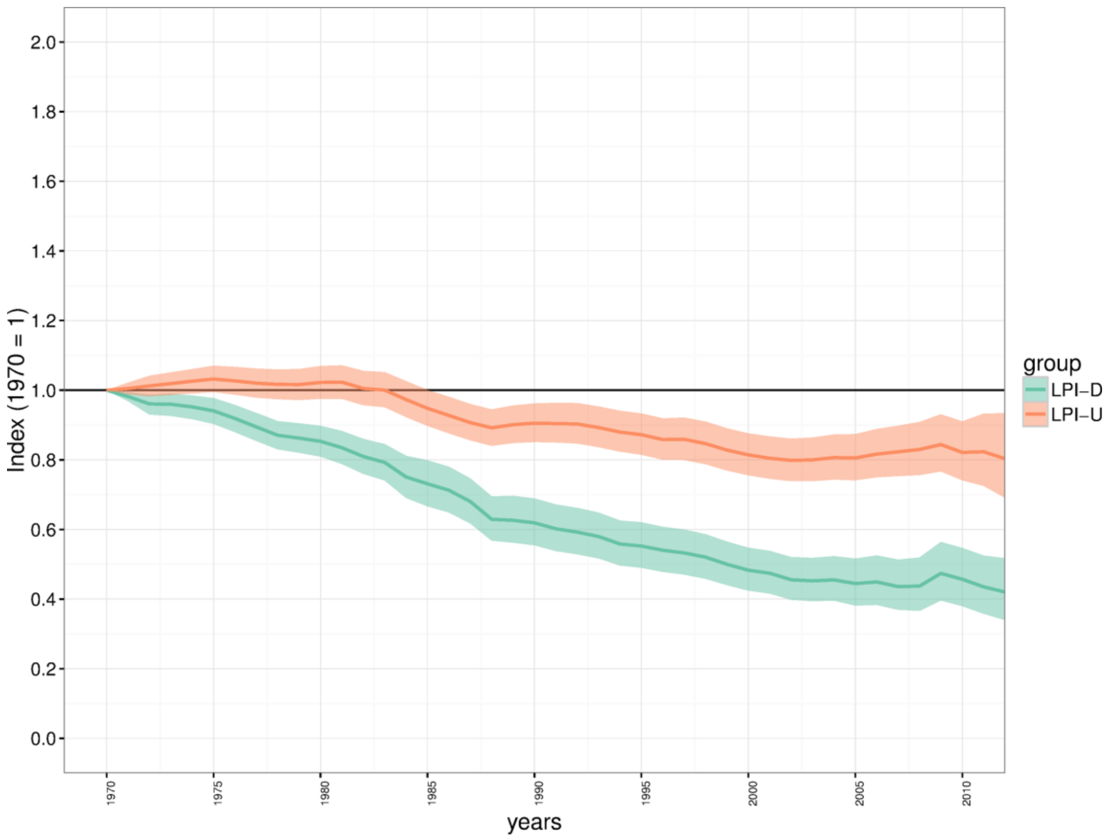

Fig. 5. The global Living Planet Index, 1970 to 2012. The red curve is unweighted, while the green curve is the biodiversity-weighted index of McRae et al. (2017). Reproduced from McRae et al. (2017).