

Special Issue: Long-term ecological research

Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time

Anne E. Magurran¹, Stephen R. Baillie², Stephen T. Buckland³, Jan McP. Dick⁴, David A. Elston⁵, E. Marian Scott⁶, Rognvald I. Smith⁴, Paul J. Somerfield⁷ and Allan D. Watt⁴

¹ Scottish Oceans Institute, School of Biology, University of St Andrews, St Andrews, Fife KY16 8LB, UK

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

³ Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews KY16 9LZ, UK

⁴ Centre for Ecology and Hydrology, Bush Estate, Penicuik, Midlothian EH26 0QB, UK

⁵ Biomathematics & Statistics Scotland, Craigiebuckler, Aberdeen AB15 8QH, UK

⁶ Department of Statistics, University of Glasgow, Glasgow G12 8QW, UK

⁷ Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, UK

The growing need for baseline data against which efforts to reduce the rate of biodiversity loss can be judged highlights the importance of long-term datasets, some of which are as old as ecology itself. We review methods of evaluating change in biodiversity at the community level using these datasets, and contrast whole-community approaches with those that combine information from different species and habitats. As all communities experience temporal turnover, one of the biggest challenges is distinguishing change that can be attributed to external factors, such as anthropogenic activities, from underlying natural change. We also discuss methodological issues, such as false alerts and modifications in design, of which users of these data sets need to be aware.

Why long-term datasets are important in biodiversity research and monitoring

We live in an era in which there is unprecedented concern about biodiversity, with 2010 the UN's International Year of Biodiversity and the target date (Convention for Biological Diversity, www.cbd.int/2010-target/) for having reduced the rate of biodiversity loss [1]. Data that can be used to monitor biodiversity, and to gauge changes in biodiversity through time, are essential. However, lack of information on the background rates and direction of change in ecological systems can make it difficult to detect the signature of anthropogenic impacts. Moreover, ecologists are increasingly aware that they have limited knowledge of temporal changes in ecological communities. For these reasons there has been an upsurge of interest in long-term data sets, some of which were initiated at the time Darwin was contemplating the evolution of biological

diversity. In this review we explore the opportunities that long-term datasets offer for tracking, and understanding, temporal changes in biodiversity at the community level and consider the challenges involved in exploiting these datasets.

Long-term datasets

Long-term datasets are as old as ecology itself [2]. In the same way that astronomy flourished once observers began to systematically document the positions of stars and planets [3], the development of ecology as a discipline is linked to the accumulation of data on the distribution and abundance of species in space and time. These data collections were often initiated to answer applied questions. For example, The Park Grass Experiment at Rothamsted in southern England, now the longest running ecological experiment [4], was founded in 1856 to examine the effect of fertilizers on yield in hay meadows [5,6]. However the ecological value of the data collected at Park Grass soon became clear and they have been used to tackle problems ranging from the evolution of adaptation at a local scale to the link between community composition and climatic perturbation [4]. Similarly, the Continuous Plankton Recorder [7] began in 1925 with the goal of mapping oceanic plankton and relating that to fisheries, but has proved invaluable in addressing many questions, including community responses to ocean warming [8]. These pioneering studies proved the worth of long-term datasets and are part of a growing number of temporally extended studies including the North American breeding bird survey [9], the UK's Environmental Change Network [10], the International Bottom Trawl Survey [11], the Wageningen [12] and Cedar Creek [13] experimental plots (the latter being one of the sites in the US Long Term Ecological Research

Corresponding author: Magurran, A.E. (aem1@st-andrews.ac.uk).

Network, www.lternet.edu) and the National Ecological Observatory Network (NEON, www.neoninc.org) (see [Supplementary Material for further information and examples](#)). However, although growing awareness about ecosystem change has underlined the importance of long-term datasets, there are still relatively few biodiversity time-series that span decades [14] or include tropical and southern temperate localities. Because the drivers of biodiversity can vary geographically [15,16], researchers cannot necessarily assume that responses to change documented in long-term datasets will be universal, even where the same types of organisms are involved.

A long-term dataset in the context of biodiversity research is simply information on the variety, and ideally the abundance, of species (or other taxonomic units) at one or more locations at a number of points in time. For the purposes of this paper we consider biodiversity as relating to a group of organisms that form an ecologically coherent unit, such as fish in a pond, vegetation in an alpine meadow, arthropods in an experimental plot, or birds in a landscape. Although we are primarily interested in biodiversity at the community level, we recognize that long-term datasets can also be used to examine temporal patterns at other hierarchical levels including genes, populations and ecosystems. The duration, frequency of sampling, spatial extent and taxonomic focus of these time series can vary markedly from study to study, or even within the same study. Indeed there are few long-term datasets where there has not been variation in sampling methodology, intensity and interval, often as the result of the waxing and waning of research priorities [14]. These factors mean that it can be difficult to make comparisons between studies, or to draw robust conclusions from a single data set.

There are two key questions concerning the temporal component of biodiversity: (1) what is the underlying level of temporal turnover in a community, and (2) is biodiversity changing relative to this background turnover, for instance due to an identified anthropogenic impact or in response to an experimental treatment? We discuss each of these and highlight methodological issues that should be taken into account when analyzing long-term datasets.

(1) Temporal turnover

Although it was clear to 19th century researchers, including Darwin [17], that ecological communities are not static, and temporal turnover underpins key ecological principles such as MacArthur and Wilson's theory of island biogeography [18,19], it is not always appreciated that species lists grow when samples are taken at successive points in time, mirroring the form of species–area curves [20–23]. This is true not just of communities undergoing succession or some other form of directional change, but also of mature communities that appear to be in a steady state. Species–time curves have implications for the analysis of long-term data sets. First, the exponent of the power curve in a plot of species against time can be used as a measure of turnover. This value usually falls within 0.2–0.4, with lower values associated with greater richness [24]. Second, given the non-linear relationship of the species–time curve one cannot deduce the species richness of a shorter (or longer) time period by simple

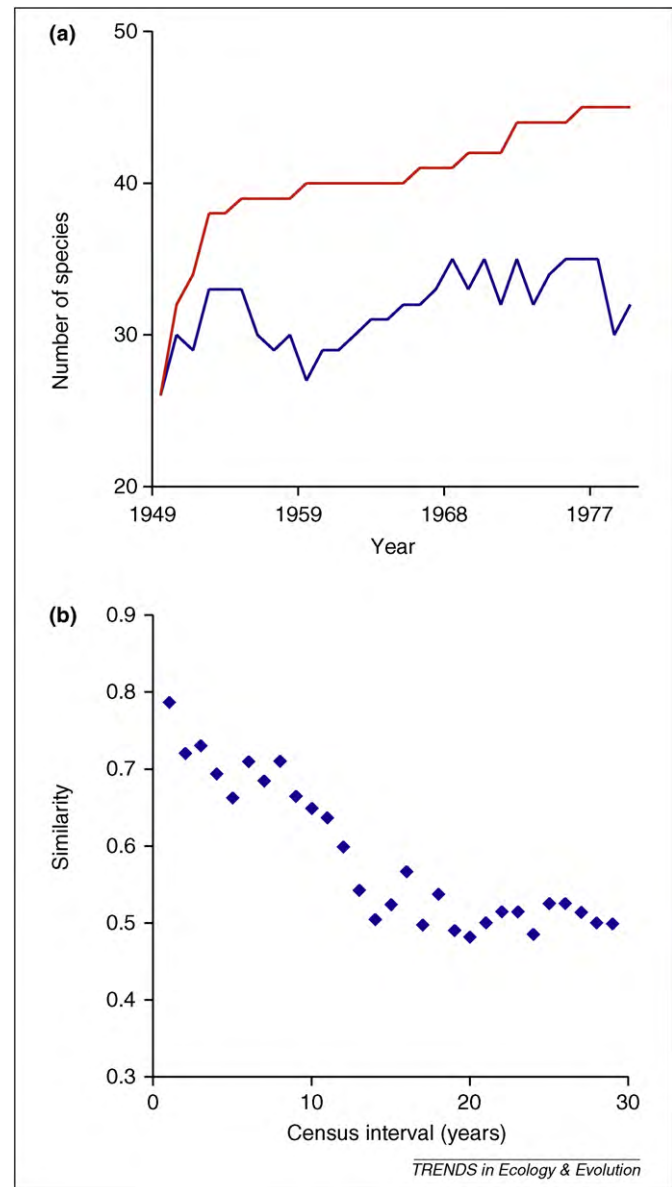


Figure 1. (a) The accumulation of species through time in a bird community in a British woodland (upper line) in relation to the number of species recorded in each year of sampling. (b) The turnover in this community (measured as Bray–Curtis similarity) in relation to census interval. The graph plots the similarity between the first and subsequent years in the time series. The Figure is based on data from the Eastern Wood bird community [76].

division or multiplication (Figure 1). Typically there will be around twice as many species detected in a decade as in a single year [22]. Rates of temporal turnover vary amongst ecosystem types [25] and in relation to local environmental factors [26], with variable responses to the same disturbance events [27]. This makes it difficult to predict the extent of turnover likely given certain types of change.

Species–time curves are only one way of assessing turnover. Alternative approaches include the use of similarity and distance measures (e.g. [28]), tailored turnover indexes (e.g. [29]) and methods that track labelled species [30] or species ranks [31]. When census intervals vary, as they do in many long-term datasets, estimates of turnover made using these methods can be biased [23] (Figure 1). It is possible to make corrections that draw on species–time curves, or that infer the number of unseen species [32].

However, few species abundance models incorporate temporal turnover [23,33], making it harder to take account of shifting species abundances. Another complication, particularly when sampling changes, is that estimates of turnover can be affected by variation in the detectability of species or individuals [34].

(2) Measuring change in biodiversity

Assessing change using diversity indexes

In principle any measure of diversity (e.g. species richness, a diversity statistic such as the Shannon or Simpson index, heterozygosity, and an index of functional or trait diversity [35]) can be plotted against time, and examined for trend [36]. A trend can take the form of either an abrupt change or a gradual change (which might or might not be linear). Examining the pattern of change in diversity statistics has the advantage of being straightforward and flexible; researchers can explore trends in different components of diversity (such as species or genes) or compare patterns across taxa. There have been few studies that have examined temporal patterns using measures of diversity other than species richness or synoptic statistics like the Shannon index. Examples that have used alternative measures include analyses of temporal turnover in the phylogenetic diversity of a Californian rockfish assemblage [37], of taxonomic distinctness of marine macrobenthos [38], and of changes in the rank–abundance and rank–energy distributions of desert rodents through time [39,40]. Diversity measures appear to be more robust to changes in sampling methods than population-level metrics (e.g. [41,42]). A critical issue is whether there is sufficient power to detect change. This is discussed in **Box 1**.

In the case of single-species surveys, change is often quantified by fitting a time-series model, so that autocorrelation is accounted for [43], and the case for applying such methods to community measures of biodiversity can be equally strong. We can go further and embed population dynamics models in inference, so that the autocorrelation is accounted for through the demographic processes (e.g. birth, death and movement rates), which themselves can be modelled using environmental covariates and/or random effects [44]. For long-lived organisms, these demographic rates might exhibit less correlation across time than population size, and so might be more sensitive indicators of changes in population performance than population size itself. As surveys become more data-rich and ambitious in scope, we can anticipate similar developments for modelling multi-species measures of biodiversity.

Changing variance as an indicator of change

Sometimes a community change, such as the collapse of a fishery [45,46], is so dramatic that the question is less about quantifying a dramatic shift in structure than about anticipating and ideally preventing it. Despite the magnitude of the changes involved, regime shifts have proved difficult to predict. Carpenter and Brock [47] argue that increased variability in biomass and other community attributes can be indicators of change from one state to the next; mathematically, this would be expected to happen as a system approaches a state of neutral equilibrium

Box 1. Power and the detection of trends

An important aspect of designing a survey is the effort required to achieve adequate precision to identify trends. In the context of a survey to quantify a trend in biodiversity at a single site, the length of the time series and the precision of measurement at each time point are key. If quantifying trends in regional biodiversity, the number of sample plots included in the survey is also critical. Other factors that affect precision for regional biodiversity estimation include plot size and frequency of surveys, and the natural variability in abundance of species.

Typically, power calculations are conducted to assess the level of effort required. If we consider a null hypothesis of no trend in a biodiversity measure, we can estimate the probability that this null hypothesis will be rejected at some significance level for a range of values for the true trend. We would like this probability to be high, except when the true trend is close to zero. We choose a sample size (number of plots) so that this probability is acceptably high (say 0.9) when the true trend is as specified. Often we assume the alternative hypothesis consists of a trend that is linear on either the scale of the observations or on a log scale, which corresponds to an exponential trend on the untransformed scale as would be expected with a fixed percentage change. For long time-series, the assumption of a linear trend can be useful for power calculations, but is generally implausible in practice. When analysing such data we might seek to estimate smooth, non-linear trends. In this circumstance, power calculations are more difficult because the alternative hypothesis is not readily specified. It is conceptually simpler and often more useful then to base the design on confidence interval length for some quantity of interest, such as percent change over a five-year period. We then estimate how many plots we need to sample to give a confidence interval of a specified length.

We would normally use data from a pilot survey to allow us to estimate the variability of future estimates of trend (conditional on some model of random variation) and hence determine sample size. If the pilot survey is limited, e.g. due to financial constraints, allowance might need to be made for uncertainty in the estimation of variability [70]. Nonetheless, pilot surveys can save substantial amounts of money in the longer term by ensuring that the survey design is sufficient for the purpose for which it is intended.

with reduced deterministic pull back towards a stable equilibrium, so in theory can be used as a warning signal even in systems with unknown dynamics. However, although these indicators do seem to provide a strong indication that a system is reaching a tipping point, they might not always provide enough warning to avert a regime shift, particularly where substantial policy changes are required [48]. In this context it will be of practical importance, but may be statistically impossible, to distinguish direct anthropogenic change which responds to policy shifts, from more diffuse and natural causes, such as an increasing variability in the weather.

Headline indicators and trends

Headline indicators and alerts are also used to identify change. In contrast to the analyses just discussed, which take all community data together and test for patterns, they build up a picture by accumulating information from different taxa.

Headline indicators such as the Living Planet Index (LPI) [49] are popular and appeal to politicians and policy-makers. Perhaps because they provide a deceptively simple summary of very complex, multiple datasets, they can easily be misinterpreted.

The LPI is an indicator ‘designed to monitor the state of the world’s biodiversity’. Using a geometric mean, it

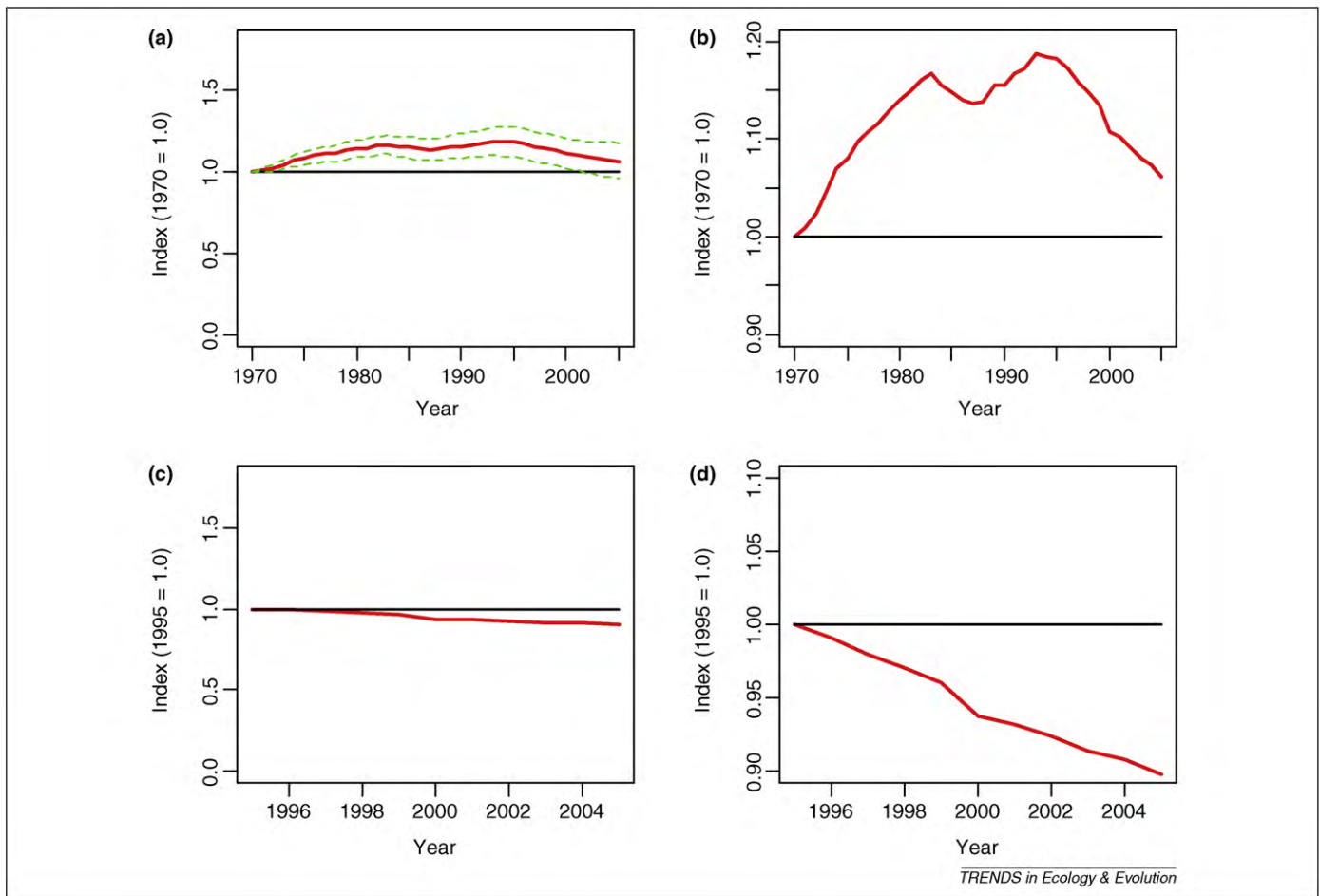


Figure 2. Living Planet Index for temperate regions of the world at various scales and starting years. The dashed lines in (a) represent pointwise 95% confidence limits. The base year is taken to be 1970 for (a) and (b) and 1995 for (c) and (d). The scale on the y-axis starts at 0.0 in (a) and (c) but is truncated in (b) and (d). The figure is based on data from www.panda.org/about_our_earth/all_publications/living_planet_report/lpr_2008/.

combines estimated trends in relative abundance of nearly 5000 populations, representing nearly 1700 species of mammal, bird, reptile, amphibian and fish. The index is arbitrarily scaled to be 1 in 1970, the baseline year. The global index is obtained by combining two indices, one tropical and one temperate. The temperate index is shown in Figure 2 with panel (a) corresponding to the published plot. The interpretation of the plot in the Living Planet Report 2008 (www.panda.org/about_our_earth/all_publications/living_planet_report/lpr_2008/) is that it shows 'little overall change' during 1970–2005. By contrast, the tropical index fell by about 50% in the same period.

There are several potential pitfalls in using this and similar indexes, and here we stress that we are using the LPI as an example of issues involved and are not singling it out for criticism. First, confidence intervals are given for the relative index, so the interval has length zero for the baseline year (1970 for the LPI), and gets progressively wider with time (see Figure 2a). Second, by changing the scale on the y-axis, trends can appear either small or large (compare Figure 2a with 2b, and 2c with 2d). Because it is difficult to interpret what, say, a 10% decline represents, it can be presented either as 'little change' or as 'dramatic decline', depending on the writer. Third, choice of base year is arbitrary, and different choices can lead the reader to very different conclusions (compare Figure 2a with 2c).

Finally it is based on a limited list of species with a bias towards vertebrates, reflecting the nature of the data that are available and so potentially introducing selection bias if trends in species included are not random with respect to the 'population' of species which the index purports to represent. It will be easier to overcome some of these pitfalls than others. For example, it would be informative to explore the consequences of using different base years, or of using a run of several years rather than a single year, to set confidence limits. However, extending indicators to include a wider range of taxa will require coordinated fieldwork with considerable additional expense for taxa not monitored otherwise.

Baselines and alerts

Alerts make assessments about the status of individual species. The IUCN's Red List (www.iucnredlist.org) is perhaps the best-known example of this approach. A community-wide perspective can be obtained by tallying the numbers or proportions of species in various categories, and by tracking the changes in status of these assemblages through time [50]. The alert system used by the British Trust for Ornithology (BTO) and other conservation organisations in the UK [51] illustrates some of the issues involved. Although not the only criterion, the primary way in which an alert is triggered is if there has been a

reduction in the size of the UK population of a bird species over a specified time period (often 25 years). If the estimated reduction is within 25–50%, an amber alert is triggered. If it is over 50%, a red alert is triggered. These statistics are calculated annually and provided on an advisory basis. More formal species lists (the Birds of Conservation Concern List) are published about every five years, grouping species into three categories: red, amber or green. These lists also incorporate expert opinion.

One difficulty of this approach is that false alerts can occur, and so any quick fixes might erroneously be considered to be effective. This is related to the speed camera phenomenon. When speed cameras are positioned at places with a high incidence of road accidents (termed ‘blackspots’ in the UK), a reduction in the accident rate is often recorded. However, this is not necessarily because the camera is having an effect. The identification of blackspots is based on accident rate. If this rate is high in a given time period purely through random fluctuation, then the rate in the following time period will tend to be lower (closer to the underlying rate) on average (an example of regression to the mean [52]), and a reduction in the accident rate will be noted whether or not a camera is placed at the location. Similarly, for scarce species, sampling fluctuation can lead to an alert being triggered, when the true reduction (if any) in the population is not of a magnitude to merit an alert. Such problems are minimized within the BTO system by using smoothing splines to remove short-term fluctuations in population trends and by taking account of the precision of the change measures within the alerts process [53]. Modelling the time-series of relative abundance indices can further reduce the problem of false alerts [54]. However it is done, the balance between sensitivity (providing early alerts of genuine change) and specificity (avoiding false alarms) needs careful consideration, as do the costs and consequences of courses of action which might ensue [55].

Another problem with the alert system is that it is heavily dependent on the baseline. The methods of Brooks *et al.* [54] ameliorate this problem by replacing the annual estimates of relative abundance by predicted relative abundances, which have greater precision and smaller fluctuations. However, there is still the difficulty that the species might have been unusually abundant or scarce in the baseline year. For example, the herring gull (*Larus argentatus*) is on the UK red list [56]. Herring gulls were especially abundant 25 years ago, a result of anthropogenic factors such as discards of fish and widespread use of open rubbish tips. Further, some species naturally cycle and whether alerts are triggered might depend on where a population was in its cycle in the baseline year. For all of these reasons it is essential that alerts are interpreted in relation to the ecology of the species being considered. While this can be relatively easy for well-studied groups such as European bird populations, it might be impossible for less well-studied groups.

Evaluating change in the rate of change

The difficulties of detecting a trend when a diversity measure or headline indicator is plotted against time can be circumvented by focussing on change in the rate

of change, as was done by the Convention on Biological Diversity, which seeks ‘a significant reduction of the current rate of biodiversity loss’ by 2010. Buckland *et al.* [36] combined smoothed trends in relative abundance using a geometric mean, and then numerically estimated the second derivative of the composite trend, to identify years over which there was a significant change in the rate of change: confidence intervals for the second derivative in each year could also be calculated. These results are unaffected if the baseline year or the scale on the y-axis is changed. Another context in which change in the rate of change has been assessed is provided by Wonham and Pachevsky [57], who used a null model to test whether there had been an increase in the rates at which invasive species were becoming established in communities.

Of course this approach does not resolve other difficulties inherent in the estimation of temporal trends, such as the inclusion of a far-from-random selection of taxa. Nevertheless, change in the rate of change is a more robust way to assess biodiversity trends than is percentage change since an arbitrary baseline year.

Synthesis: opportunities and challenges in using long-term data sets

We have demonstrated that long-term datasets are an essential resource in biodiversity research and monitoring. However, we have also shown that it is not always as easy as it first appears to determine whether the underlying nature of a community is changing, and by how much. In this final section we highlight opportunities, identify challenges and suggest solutions.

Going back to the baseline

Temporal patterns of biodiversity have received much less attention than spatial ones. Although there is a substantial body of research on directional change, such as succession, natural fluctuations in mature communities have been neglected. This means that there are considerable opportunities for basic research that will in turn inform monitoring programmes and aid decision makers. We note three areas that will repay investigation. First, despite the fact that the early developers of species abundance models, notably Fisher [58] and Preston [59,60], explicitly incorporated temporal turnover, later workers have taken little account of temporal changes in communities. Indeed time and space are often confounded [61]. A new generation of species abundance models that make predictions about how communities change through time will not only shed new light on how communities are structured and biodiversity is maintained, but will also provide improved benchmarks for monitoring. Long-term datasets will be essential for model formulation and testing. Second, there is considerable scope for investigations that partition diversity in both space and time [23,29,62–64]. This will help answer longstanding questions about the equivalence of space and time [60], and contribute to the management of biodiversity at the landscape level. Finally, we know very little about correlated responses in diversity through time, both in the sense of whether different diversity

Box 2. When schemes and objectives change

Objectives of long-term surveys typically change over time in response to changing priorities, as occurred when the 2010 Biodiversity Target of the Convention on Biological Diversity was adopted. If objectives change to a degree that undermines the value of a survey, there are broadly three options. The first is to continue without change, so that the time-series already gathered is not compromised. It is essential in this case to explore the limitations of the survey given the new objectives, and to evaluate whether it still represents value for money. The second option is to retain the design of the survey, but to gather additional data to address the revised objectives. Here, provided the data can be reduced to data comparable with those collected before the change, the existing time-series is not compromised. The third option is to replace the old survey. We illustrate this option using the changeover from the Common Birds Census (CBC, www.bto.org/birdtrends2009/cbc.htm) to the Breeding Bird Survey (BBS, www.bto.org/bbs/index.htm) in the UK in the 1990s.

The CBC was a survey of breeding birds, primarily in two habitats, woodland and farmland. Subjectively-chosen plots were surveyed by volunteers using territory mapping, and time trends in species-specific relative abundance estimated. The estimates were largely driven by trends in southern Britain, where most volunteers lived. Further, it was not possible to provide statistically sound estimates of regional trends even in southern Britain, because the plots were not selected randomly. Hence a decision was made in the early 1990s to replace it by the BBS. Plots for the BBS are 1-km squares throughout the UK, selected according to a stratified random sampling scheme. Sampling intensity within a geographic stratum is determined by the number of volunteers available. Within a plot, two parallel transects, each 1-km long, are walked twice each breeding season, and counts of each species in each of three distance intervals are made.

The problem of how to link the two time-series together was resolved by continuing both surveys in parallel during 1994–2000. Although overlap in a single year would appear to be all that is required, this leads to poor precision in alignment of the two series. Further, by having seven years of overlap, it was possible to assess whether trends estimated by the two schemes were consistent. For most species, and considering southern Britain only, they did indeed prove to be largely consistent [71], giving greater confidence in trends estimated by combining the two time-series of estimates.

metrics show the same pattern, and whether different taxa respond in the same way.

Dealing with changes in methodology

One of the biggest problems that confronts users of long-term datasets is a change in methodology part way through the study. This can arise if the aims of the investigation change, or because new statistical techniques or sampling technologies are developed, or due to external factors such as a change of land use. Sometimes managers are faced with the dilemma of whether to stick with sub-standard methodology for the sake of comparability, or whether to switch to a better design. Although the decision can be finely-balanced, change can be managed successfully as Box 2 shows.

Collecting sufficient data

Biodiversity data are time-consuming and expensive to collect [65] with the result that data sets are frequently short-term or incomplete [14]. One solution is to involve volunteers (citizen scientists) in data collection [66]; another is to make use of existing opportunities, as for example the Sir Alistair Hardy Foundation does by cooperating with the commercial ships that tow its plankton

Box 3. Citizen science

The concept of citizen science, a two-way cooperation between the scientific and the public communities in long-term monitoring programs, is not new. One of the oldest citizen science programs is organised by the National Audubon Society, namely the Christmas Bird Count (www.audubon.org/bird/cbc/index.html). In 1900, 27 observers took part in the first count in 25 places in the United States and Canada and since then the counts have been held every winter. The 101st count, in the winter of 2000–2001, involved 52 471 people in 1823 places in 17 countries. Advances in electronic recording and communication systems via the world wide web and mobile phones have led to a resurgence with projects like Project-Budburst (www.windows.ucar.edu/citizen_science/budburst/index.html) using real-time mapping with Google Maps, photo-sharing and scientist blogs to engage younger citizens. Involving unpaid volunteers has the advantages of being economic, can extend the geographic range of study sites and the frequency of visits, and can have huge educational advantages for all levels of society [72].

Successfully harnessing the extra power brought through citizen science requires adherence to statistical principles. The survey requires a proper sampling strategy, such as a stratified random scheme, and strict recording protocols that are clearly understood. It can be useful to incorporate checks such as photographic records or electronic identifications. When designing protocols, simplicity of tasks and sympathetic understanding, for example by altering the frequency of observations to fit people's lifestyles, benefits the data capture and subsequent analyses. There are likely to be site locations and times which either cannot be covered by available volunteers or are under-represented, and paid staff might be required if avoiding such omissions is important to the study. There will be variation in volunteer ability which should be modelled in the analysis; thus the UK Breeding Bird Survey uses as units of analysis distinct combinations of site and observer. Although this confounds the site and observer effects in the analysis and so requires a relatively dense site network to recover geographically useful results, it does allow for loss or change of observers over years.

It is essential to have clear scientific objectives, unambiguous data collection protocols, and regular reporting back of data summaries in order to ensure continued support and data quality through promotion of best practice.

samplers [7]. Both solutions can involve exchanging control over data collection for increasing coverage. As a result of this trade-off, additional modelling might be required. Box 3 further explores the opportunities for citizen science and the ways in which this type of data collection can be optimized for biodiversity research.

Sampling issues

A striking feature of long-term datasets is how much they vary in design, particularly with respect to sampling frequency. For example, the datasets used to evaluate trends in biodiversity in relation to the 2010 target (www.twentyten.net) include the 'extent of assorted ecosystems indicator' which covers: seagrass, for which compiled data span c. 100 years but where the frequency of assessments varies from annually, to every three years, to decadal; mangrove swamps where there are data from only four time-periods; and coral reefs where records started in 1970 but there are measurements every couple of years. Sampling interval and intensity have a significant impact on both the amount of diversity detected and the extent of change in that diversity [21]. Palaeontologists face parallel problems when evaluating origination and extinction rates in fossil communities, except that the sampling issues confronting them are much greater than those that face ecologists [67]. Comparisons

Box 4. Uncertainty in the context of biodiversity assessment

“Quantifying an idea as apparently vague as uncertainty came comparatively late to science. This is perhaps unsurprising when we acknowledge a distinction between chance or aleatory uncertainty, concerned with essentially random phenomena, and probability or epistemological uncertainty, which concerns lack of knowledge about unique and potentially verifiable events and so is essentially a measure of ignorance.” [73]

This article has focused on the role of statistics derived from long-term datasets to describe biodiversity change. Their scientific uncertainty is assessed through combining several known sources of statistical error, some discussed here, including measurement precision, sampling procedures and basic ecological processes such as species turnover. While this allows the scientific questions to be answered, environmental policymakers reach decisions based on expert knowledge, existing research and statistics, and stakeholder consultation. Evidence is fundamentally uncertain, and a key question is how best to acknowledge, quantify and respond, not only to scientific uncertainty but also to ignorance, a state of not knowing from which springs both scientific discoveries and unpleasant ‘surprises’ [74]. Increasingly the topic of communicating uncertainty is high on the science/policy agenda, e.g. the Royal Society Discussion meeting: ‘Handling Uncertainty in Science’ in March 2010 (royalsociety.org/Handling-uncertainty-in-science/).

One science area where there has been one of the most visible discussions of uncertainty is in the field of climate change science, and Intergovernmental Panel on Climate Change (IPCC) have for recent reports developed guidelines that can have relevance to other science areas, including biodiversity. The IPCC guidance on description of findings is especially pertinent [75]. If there is the power to identify a trend or change, the outcome might be reported as an increase, a decrease, or no significant change. They advise that it is important to explain the basis for the conclusion reached and to set out the extent to which alternative outcomes would not have been expected. Changes that have a reasonable likelihood of occurring should be included even where they are not certain.

With regard to biodiversity indicators (or any other indicator for that matter), we have suggested that an important first step (though not the only one) is to quantify the uncertainty arising from the statistical estimation (directly linked to the observed variability); this is close to the notion of structural uncertainty used by the IPCC [75]. Surprisingly, this has not commonly been done, although there is widespread recognition of variation in the natural world. There are many other sources of uncertainty (including unpredictability and value uncertainty [75]) and transparency in reporting, including a calibrated narrative scale of probability and likelihood, is essential. Our recommendations suggest a small, but entirely practicable, step forward.

amongst long-term data sets, and assessments that combine data sets, need to take account of sampling, for example by including census interval in models.

Trends in biodiversity: distinguishing anthropogenic change from background change

Barely a day passes without a news story about the decline in biodiversity. However, the fact that ecological communities constantly experience temporal turnover, and that consequently some species will not only fluctuate markedly but also become either locally or globally extinct, is something that, while well appreciated by ecologists generally, is often omitted from popular news stories. Multiple factors influence species abundance and distribution [68,69], and since most ecological communities are already impacted by habitat loss and harvesting, it can be difficult to make the distinction between natural and anthropogenic change. Conveying this complexity to a non-specialist audience is

not easy as it requires a wider appreciation of the natural variability of communities. Reports of biodiversity loss attributed to an impact (such as a new type of fisheries policy or climate change) should be accompanied by uncertainty values that take account of baseline change. The uncertainty should also include measures of precision on the estimates of change, to minimize the risk that policy advice is based on an estimated change that is not real (see Box 4 for further discussion of this point, as well as the issues raised when communicating uncertainty).

Prospectus

The earliest ecologists showed great foresight when they initiated ecological experiments and began systematic data collections. Without these data sets we would have a more limited understanding of how ecological communities function, and be less well-equipped to detect the losses of biodiversity that have resulted from the anthropogenic changes that have accompanied the expansion of the human population and its associated activities. Long-term datasets can be used for multiple purposes, and in many cases have helped answer questions that their founders never considered. As we have shown there are important methodological issues to be considered, particularly when drawing conclusions about the extent of change in ecological communities. We expect long-term datasets to become an increasingly important resource in biodiversity research, to have growing application in the monitoring and conservation of biological diversity, and to be used to answer questions that neither we, nor the originators of the surveys, will have anticipated.

Acknowledgements

AEM acknowledges the Royal Society of Edinburgh for a support fellowship. DAE was funded by the Scottish Government Rural and Environment Research and Analysis Directorate. PJS acknowledges support from NERC's Oceans 2025 programme and Defra (ME3109). The authors acknowledge the support of the Royal Statistical Society through their funding of the Panel on Statistics for Ecosystem Change, and the referees for their helpful comments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.tree.2010.06.016](https://doi.org/10.1016/j.tree.2010.06.016).

References

- 1 Butchart, S.H.M. *et al.* (2010) Global biodiversity: indicators of recent declines. *Science* 328, 1164–1168
- 2 McLean, A.R. and May, R.M. (2007) Introduction, In *Theoretical Ecology* (3rd ed.) (May, R.M. and McLean, A.R., eds), pp. 1–6, Oxford University Press
- 3 Holmes, R. (2009) *The Age of Wonder*, Harper Press
- 4 Silvertown, J. *et al.* (2006) The Park Grass Experiment 1856–2006: its contribution to ecology. *J. Ecol.* 94, 801–814
- 5 Lawes, J. and Gilbert, J. (1880) Agricultural, botanical and chemical results of experiments on the mixed herbage of permanent grassland, conducted for many years in succession on the same land. *Phil. Trans. R. Soc. B.* 171, 189–416
- 6 Lawes, J. *et al.* (1882) Agricultural, botanical and chemical results of experiments on the mixed herbage of permanent meadow, conducted for more than twenty years on the same land. II. The botanical results. *Phil. Trans. R. Soc. B.* 173, 1181–1413
- 7 Richardson, A.J. *et al.* (2006) Using continuous plankton recorder data. *Prog. Oceanogr.* 68, 27–74

- 8 Kirby, R.R. *et al.* (2007) Climate effects and benthic-pelagic coupling in the North Sea. *Mar. Ecol. Prog. Ser.* 330, 31–38
- 9 Sauer, J.R. *et al.* (2008) *The North American Breeding Bird Survey, Results and Analysis 1966–2007*, USGS Patuxent Wildlife Research Center
- 10 Morecroft, M.D. *et al.* (2009) The UK Environmental Change Network: emerging trends in terrestrial biodiversity and the physical environment. *Biol. Conserv.* 142, 2814–2832
- 11 Rijnsdorp, A.D. *et al.* (1996) Changes in the abundance of demersal fish species in the North Sea between 1906–1909 and 1990–1995. *ICES J. Mar. Sci.* 53, 1054–1062
- 12 Bezemer, T.J. and van der Putten, W.H. (2007) Ecology: diversity and stability in plant communities. *Nature* 446, E6–E7
- 13 Tilman, D. *et al.* (2006) Biodiversity and ecosystem stability in a decade-long experimental grassland. *Nature* 441, 629–632
- 14 Wolfe, D.A. *et al.* (1987) Long-term biological data sets: their role in research, monitoring, and management of estuarine and coastal marine systems. *Estuaries* 10, 181–193
- 15 Hawkins, B.A. *et al.* (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84, 3105–3117
- 16 Chown, S.L. *et al.* (2004) Hemispheric asymmetries in biodiversity: A serious matter for ecology. *PLoS Biol.* 2 DOI: 10.1371/journal.pbio.0020406
- 17 Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, John Murray
- 18 MacArthur, R.H. and Wilson, E.O. (1967) *The Theory of Island Biogeography*, Princeton University Press
- 19 Rosenzweig, M.L. (1995) *Species Diversity in Space and Time*, Cambridge University Press
- 20 van der Gast, C.J. *et al.* (2008) Temporal scaling of bacterial taxa is influenced by both stochastic and deterministic ecological factors. *Environ. Microbiol.* 10, 1411–1418
- 21 White, E.P. (2007) Spatiotemporal scaling of species richness: patterns, processes, and implications. In *Scaling biodiversity* (Storch, D. *et al.*, eds), pp. 325–346, Cambridge University Press
- 22 White, E.P. *et al.* (2006) A comparison of the species-time relationship across ecosystems and taxonomic groups. *Oikos* 112, 185–195
- 23 Magurran, A.E. (2010) Measuring biological diversity in time (and space). In *Biological Diversity: Frontiers in Measurement and Assessment* (Magurran, A.E. and McGill, B.J., eds), Oxford University Press
- 24 Shurin, J.B. (2007) How is diversity related to species turnover through time? *Oikos* 957–965
- 25 Korhonen, J.J. *et al.* (2010) A quantitative analysis of temporal turnover in aquatic assemblages across ecosystems. *Ecology* 508–517
- 26 Werner, E.E. *et al.* (2007) Turnover in an amphibian metacommunity: the role of local and regional factors. *Oikos* 116, 1713–1725
- 27 Svensson, J.R. *et al.* (2009) Equal rates of disturbance cause different patterns of diversity. *Ecology* 90, 496–505
- 28 Chytrý, M. *et al.* (2009) Species richness and species turnover in a successional heathland. *Appl. Veg. Sci.* 4, 89–96
- 29 Jost, L. *et al.* (2010) Compositional similarity and beta diversity. In *Biological Diversity: Frontiers in Measurement and Assessment* (Magurran, A.E. and McGill, B.J., eds), Oxford University Press. (in press)
- 30 MacNally, R. (2007) Use of the abundance spectrum and relative-abundance distributions to analyze assemblage change in massively altered landscapes. *Am. Nat.* 170, 319–330
- 31 Collins, S.L. *et al.* (2008) Rank clocks and plant community dynamics. *Ecology* 89, 3534–3541
- 32 Chao, A. *et al.* (2006) Abundance-based similarity indices and their estimation where there are unseen species in samples. *Biometrics* 62, 361–371
- 33 McGill, B.J. *et al.* (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol. Lett.* 10, 995–1015
- 34 Buckland, S.T. *et al.* (2010) Biodiversity monitoring: the relevance of detectability. In *Biological Diversity: Frontiers in Measurement and Assessment* (Magurran, A.E. and McGill, B.J., eds), Oxford University Press. (in press)
- 35 Magurran, A.E. and McGill, B.J. (2010) *Biological Diversity: Frontiers in Measurement and Assessment*, Oxford University Press
- 36 Buckland, S.T. *et al.* (2005) Monitoring change in biodiversity through composite indices. *Phil. Trans. R. Soc. B.* 243–254
- 37 Pavoine, S. *et al.* (2009) Hierarchical partitioning of evolutionary and ecological patterns in the organization of phylogenetically-structured species assemblages: application to rockfish (genus: *Sebastes*) in the Southern California Bight. *Ecol. Lett.* 12, 898–908
- 38 Warwick, R.M. and Clarke, K.R. (1995) New 'biodiversity' measures reveal a decrease in taxonomic distinctness with increasing stress. *Mar. Ecol. Prog. Ser.* 129, 301–305
- 39 White, E.P. *et al.* (2004) Trade offs in community properties through time in a desert rodent community. *Am. Nat.* 164, 670–676
- 40 Thibault, K.M. *et al.* (2004) Temporal dynamics in the structure and composition of a desert rodent community. *Ecology* 85, 2649–2655
- 41 Trenkel, V.M. and Rochet, M.-J. (2003) Performance of indicators derived from abundance estimates for detecting the impact of fishing on a fish community. *Can. J. Fish. Aquat. Sci.* 60, 67–85
- 42 Wantiez, L. (1996) Comparison of fish assemblages sampled by a shrimp trawl and a fish trawl in St Vincent Bay, New Caledonia. *J. Mar. Biol. Assoc. UK* 76, 759–775
- 43 Thomas, L. *et al.* (2004) Temporal inferences from distance sampling surveys. In *Advanced Distance Sampling* (Buckland, S.T. *et al.*, eds), Oxford University Press
- 44 Buckland, S.T. *et al.* (2007) Embedding population dynamics models in inference. *Stat. Sci.* 22, 44–58
- 45 Jackson, J.B.C. *et al.* (2001) Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293, 629–637
- 46 Scheffer, M. *et al.* (2001) Catastrophic shifts in ecosystems. *Nature* 413, 591–596
- 47 Carpenter, S.R. and Brock, W.A. (2006) Rising variance: a leading indicator of ecological transition. *Ecol. Lett.* 9, 311–318
- 48 Biggs, R. *et al.* (2009) Turning back from the brink: detecting an impending regime shift in time to avert it. *Proc. Natl. Acad. Sci. U. S. A.* 106, 826–831
- 49 Loh, J. *et al.* (2005) The Living Planet Index: using species population time series to track trends in biodiversity. *Phil. Trans. R. Soc. B.* 360, 289–295
- 50 Stuart, S.N. *et al.* (2004) Status and trends of amphibian declines and extinctions worldwide. *Science* 306, 1783
- 51 Gregory, R.D. *et al.* (1999) *National and Site Based Alert Systems for UK Birds*. Research report 226, British Trust for Ornithology
- 52 Christie, S.M. *et al.* (2003) Are mobile speed cameras effective? A controlled before and after study. *Inj. Prev.* 9, 302–306
- 53 Baillie, S.R. and Rehfisch, M.M., eds (2006) *National and Site-based Alert Systems for UK Birds*. Research report 226, British Trust for Ornithology
- 54 Brooks, S.P. *et al.* (2008) Quantifying conservation concern – Bayesian statistics and the Red Data Lists. *Biol. Conserv.* 141, 1436–1441
- 55 Field, S.A. *et al.* (2005) Optimizing allocation of monitoring effort under economic and observational constraints. *Journal Information* 69
- 56 Eaton, M.A. *et al.* (2009) The State of the UK's Birds 2008. RSPB, BTO, WWT, CCW, NIEA, JNCC, NE and SNH
- 57 Wonham, M.J. and Pachevsky, E. (2006) A null model of temporal trends in biological invasion records. *Ecol. Lett.* 9, 663–672
- 58 Fisher, R.A. *et al.* (1943) The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* 12, 42–58
- 59 Preston, F.W. (1948) The commonness, and rarity, of species. *Ecology* 29, 254–283
- 60 Preston, F.W. (1960) Time and space and the variation of species. *Ecology* 41, 612–627
- 61 Magurran, A.E. (2007) Species abundance distributions over time. *Ecol. Lett.* 10, 347–354
- 62 Jost, L. (2007) Partitioning diversity into independent alpha and beta components. *Ecology* 88, 2427–2439
- 63 Sobek, S. *et al.* (2009) Spatiotemporal changes of beetle communities across a tree diversity gradient. *Divers. Distrib.* 15, 660–670
- 64 Tomašových, A. and Kidwell, S.M. (2009) Fidelity of variation in species composition and diversity partitioning by death assemblages: time-averaging transfers diversity from beta to alpha levels. *Paleobiology* 35, 94–118
- 65 Gardner, T.A. *et al.* (2007) The cost-effectiveness of biodiversity surveys in tropical forests. *Ecol. Lett.* 11, 139–150

- 66 Braschler, B. *et al.* (2010) Realizing a synergy between research and education: how participation in ant monitoring helps raise biodiversity awareness in a resource-poor country. *J. Insect Conserv.* 14, 19–30
- 67 Wagner, P.J. and Lyons, S.K. (2010) Examining extinction in the fossil record. In *Biological Diversity: Frontiers in Measurement and Extinction* (Magurran, A.E. and McGill, B.J., eds), Oxford University Press
- 68 Tolley, K.A. *et al.* (2007) Deconstructing a controversial local range expansion: conservation biogeography of the painted reed frog (*Hyperolius marmoratus*) in South Africa. *Divers. Distrib.* 14, 400–411
- 69 Walther, G.R. *et al.* (2009) Alien species in a warmer world: risks and opportunities. *Trends Ecol. Evol.* 24, 686–693
- 70 Sims, M. *et al.* (2007) Incorporating variance uncertainty into a power analysis of monitoring designs. *J. Ag. Biol. Envir. St.* 12, 236–249
- 71 Freeman, S.N. *et al.* (2007) Modelling population changes using data from different surveys: the Common Birds Census and the Breeding Bird Survey. *Bird Study* 54, 61–72
- 72 Bonney, R. *et al.* (2009) Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience* 59, 977–984
- 73 Spiegelhalter, D.J. (2008) Understanding uncertainty. *Ann. Fam. Med.* 6, 196
- 74 Harremois, P. *et al.* (2001) *Late Lessons from Early Warnings: The Precautionary Principle 1896–2000*, Office for Official Publications of the European Communities
- 75 IPCC (2005) AR4 Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties
- 76 Gaston, K.J. and Blackburn, T.M. (2000) *Macroecology*, Blackwell Science