Review of UK biodiversity indicators that provide status and trends for species

UK & England Biodiversity Indicators Quality Assurance Science Panel

Professor Ken Norris, Institute of Zoology (Panel Chair)
Professor Stephen Buckland, University of St Andrews
Professor Rhys Green, University of Cambridge
Professor Helen Roy, Centre for Ecology & Hydrology
Dr Phil Stephens, University of Durham

January 2016

Table of Contents

Introduction 3-4

Structured Indicator Reviews

B6 – Pressures from invasive species 5-7

C3b – Status of UK species of European importance 8-12

C4a – Status of priority species – relative abundance 13-18

C4b – Status of priority species – frequency of occurrence (insects) 19-23

C5a-d – Farmland birds, woodland birds, wetland birds and seabirds 24-35

C5e – Wintering waterbirds 36-38

C6 – Butterflies 39-45

C8a – Bats 46-48

D1c – Status of pollinating insects 49-50

General framework of good practice 51-55

INTRODUCTION

The quality assurance panel was convened to provide advice on improvements that could be made to a suite of UK biodiversity indicators that are used to assess the status and trends of a range of species, and which are linked in turn to a range of biodiversity targets. To do this, we conducted structured reviews of the indicators in Table 1.

Table 1. Indicators included in our structured reviews.

Indicator/measure	***UK Biodiversity Framework* measure**	**Link to Indicator, Data and Technical Paper**	***Biodiversity 2020* measure (England)**	**Link to Indicator**
Pressure from invasive species (freshwater, coastal, terrestrial)	B6a-c (GB data presented)	Indicator Data sheet Technical doc	20 (GB data presented)	Indicator
Status of UK species of European Importance	C3b	Indicator Data sheet Technical doc	4b	Indicator
Status of priority species – relative abundance	C4a	Indicator Data sheet Technical doc	4a (UK data presented)	Indicator
Status of priority species – frequency of occurrence (insects)	C4b	Indicator Data sheet Technical doc		
Farmland birds	C5a		5	Indicator
Woodland birds	C5b	Indicator	6	Indicator
Wetland birds	C5c	Data sheet	7	Indicator
Seabirds	C5d	Technical doc	8	Indicator
Wintering waterbirds	C5e		7	Indicator
Butterflies	C6a C6b	Indicator Data sheet Technical doc	5 (farmland) 6 (woodland)	5. Farmland Indicator 6. Woodland indicator
Plant Diversity	C7	No indicator at present	5 and 6	No measure
Bats	C8a	Indicator	5	Indicator

		Data sheet Technical doc[1]		
Status of pollinating insects	D1c	Indicator Data sheet Technical doc	10 (UK data presented)	Indicator

We used a structured review process to critically examine each indicator and to produce suggestions for improvements in each. We adopted the following general structure, which has been adapted to each indicator as necessary:

- Background
  - A short description of the indicator and associated trends
- Data Quality
  - Survey design
  - Where are the surveys done?
  - Changes over time in survey locations
  - Fieldwork methods
  - Data quality assurance
- Rigour of analytical methods
- Precision and bias
- Interpretation
- Conclusions
  - Improvements to data collection
  - Improvements to analysis methods
  - Improvements to interpretation

Each structured review was led by one expert but discussed by all panel members to ensure a consistent approach. A structured review of each indicator is presented in the following pages. Following these reviews, we present a synthesis of the general issues that emerged from the reviews within a good practice framework, and make some suggestions about improvements that may be made.

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

STRUCTURED INDICATOR REVIEWS

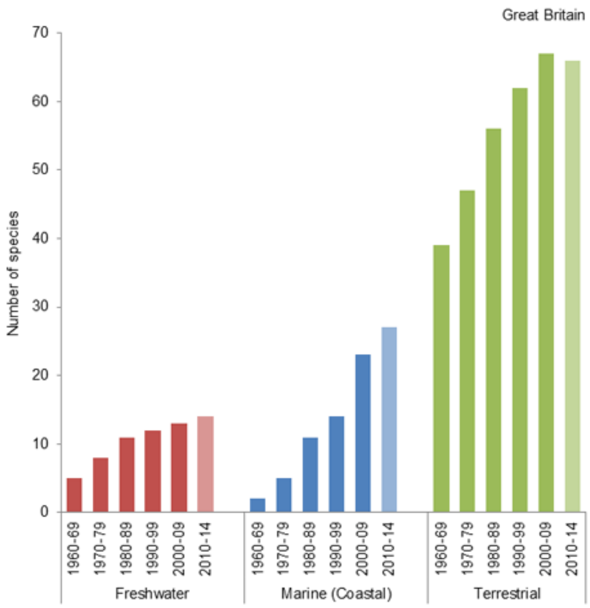
B6 - Pressure from invasive species

Helen Roy

Background

There are approximately 2000 non-native species established within Great Britain. Approximately 10-15 % of these are considered to have adverse effects on biodiversity, the economy or society and these are termed invasive non-native species (INNS). The indicator represents changes in the number of invasive non-native species (INNS) established across Great Britain. 179 INNS (38 freshwater, 34 marine and 107 terrestrial species) are included within the indicator. The species were selected by expert review of the information on impacts of INNS compiled within the GB Non-Native Species Information Portal (described by Roy et al. 2014). Occurrence data for the 179 INNS are from the NBN Gateway and used to calculate area of extent which is subsequently subject to expert validation.

Number of INNS established in or along more than 10 per cent of Great Britain's land area or coastline, 1960 to 2014



Notes: The last time period covers a shorter period than the other bars (2010–2014).
Source: Botanical Society of Britain and Ireland, British Trust for Ornithology, Centre for Ecology & Hydrology, Marine Biological Association, National Biodiversity Network Gateway.

Data Quality

Survey design

The occurrence data for the breadth of species (across all environments) included within the indicator is mostly collected by volunteer recorders associated with national schemes and societies. The main data providers are the Marine Biological Association, Botanical Society for Britain and Ireland, the British Trust for Ornithology and national schemes and societies. Data were downloaded from the NBN Gateway at 10km resolution. The area of extent was calculated using a method based on Stroh et al. 2014. The area of extent enabled species to be classified in five extent categories (defined within the indicator):

0 - Absent
1 - Not or scarcely established
2 – Established but still generally absent or at most occasional
3 – Established and frequent in part of the territory
4 - Widespread

Experts from Marine Biological Association, Botanical Society for Britain and Ireland and the British Trust for Ornithology reviewed the area of extent and the associated classification of extent. The expert review was particularly important to offset bias for species which are known to be more widely distributed than reflected in the NBN Gateway.

Data quality assurance

Occurrence data is compiled and checked by experts. The validation of the area of extent is mainly to offset bias as a consequence of lags in data flow to the NBN Gateway.

Rigour of Analytical Method

The indicator is very simple presenting number of INNS within different extent categories as figures. Statistical analysis on trends is missing and could provide a useful addition.

Precision and Bias

Observational, reporting and detection biases are inherent in occurrence data gathered through biological recording (Isaac and Powney, 2015); the distribution and activity of volunteer recorders will influence spatial and temporal coverage of the occurrence records. However calculation of the area of extent coupled with validation by experts should partially overcome the bias.

Derivation of the list of INNS is based on various evidence sources. For some (possibly many) species empirical sources of information on evidence of impact is lacking.

Interpretation

The indicator highlights changes to the extent of distribution of INNS and additionally provides a list of NNS considered to impact biodiversity and hence termed invasive. In a broad sense it provides an overview of the change in pressure from INNS.

The aim of the indicator is to fulfil obligations, as a result of the adoption of the Strategic Plan for Biodiversity (including the Aichi Targets) at the 10th Conference of Parties of the Convention on Biological Diversity, to report progress against Aichi Target 9:

Target 9: By 2020, invasive alien species and pathways are identified and prioritized, priority species are controlled or eradicated and measures are in place to manage pathways to prevent their introduction and establishment.

The indicator also fulfils reporting requirements within the GB Invasive Non-Native Species Strategy:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/455526/gb-non-native-species-strategy-pb14324.pdf

The Regulation 1143/2014 on INNS is also relevant:

http://ec.europa.eu/environment/nature/invasivealien/index_en.htm

Conclusions

We conclude that the quality, clarity and relevance of this indicator could be improved in the following ways:

Improvements to data collection

Assessment of the list of INNS on an annual basis is critical to ensure inclusion of new arrivals but also to update existing non-native species based on emerging available evidence of impacts. The focus for monitoring could be the list of INNS of EU concern which will be published in the next year.

Improvements to analysis methods

Currently the temporal trends are not analysed statistically. It is possible that statistical analysis could be applied to derive changes in extent of occurrence of INNS but the rigour of such methods would be dependent on improvements in the flow of occurrence data to ensure that the data available reflected the distribution of the species. Modelling techniques such as though used for other indicators based on occurrence data could then be employed for example occupancy modelling. However, the current status of data available would limit application of such techniques

Improvements to interpretation

The Regulation 1143/2014 on INNS entered into force on 1 January 2015 and should be considered in future developments. It would be particularly pertinent to report on the list of invasive alien (=non-native) species of EU concern within the Regulation. Interception data would be highly relevant but difficult to compile. Analysis of pathway information would also align with the Regulation.

References

Isaac, N. J., Pocock, M. J. (2015) Bias and information in biological records..Biological Journal of the Linnean Society.115, 522-531

Roy, H.E., Preston, C.D., Harrower, C.A., Rorke, S.L., Noble, D., Sewell, J., Walker, K. Marchant, J., Seeley, B., Bishop, J., Jukes, A., Musgrove, A., Pearman D. (2014) GB Non-native Species Information Portal: documenting the arrival of non-native species in Britain. Biological Invasions 16, 2495-2505

Stroh, P.A., Leach, S.J., August, T.A., Walker, K.J., Pearman, D.A., Rumsey, F.J., Harrower, C.A., Fay, M.F., Martin, J.P., Pankhurst, T., Preston, C.D., Taylor, I. (2014) A vascular plant red list for England. Botanical Society of Britain and Ireland, 184pp.

C3b - Status of UK species of European Importance

Phil Stephens

Background

This indicator (Fig. 1) is a multi-species indicator intended to summarise the conservation status of species listed on the Habitats Directive Annexes (II, IV & V). There are 125 species listed but only 93 of them regularly occur in the UK (or her waters). Consequently, the index is based on those 93, which include 36 mammal, 13 fish, 4 amphibian, 3 reptile, 16 invertebrate and 21 plant species; 16 species are marine (two marine algae, one turtle, two seals, 11 cetaceans). The index itself summarises a range of issues affecting the species, including trends in their ranges (with respect to a favourable reference range); trends in their populations (with respect to a favourable reference population); the area, quality and trends in availability of habitat; and threats to the species. This information is combined according to the table shown in Fig. 2, in order to produce the index.

Article 17 of the EU Habitats Directive requires that Member States report every six years about the progress made with the implementation of the Habitats Directive and, specifically, with maintaining and/or restoring a favourable conservation status for habitat types & species of community interest. Article 17 is prescribed by the EU and the process within the UK has its own governance and review process. Thus, the review of this indicator is different to reviews of the other species-based indicators. Here, it is required only that the panel considers how the indicator is produced and analysed, and its interpretation (i.e. not data collection). However, without considering how favourable reference values were set and data collection approaches, it is difficult to comment on credibility of the indicator. Two questions were identified for further investigation to take forward the review of this indicator: 1) Are those who assign favourable conservation status to individual species using the same approach – i.e. how clear is the methodology; how well defined are the categories, and how open to interpretation are they? 2) Is uncertainty assigned in a sensible way?

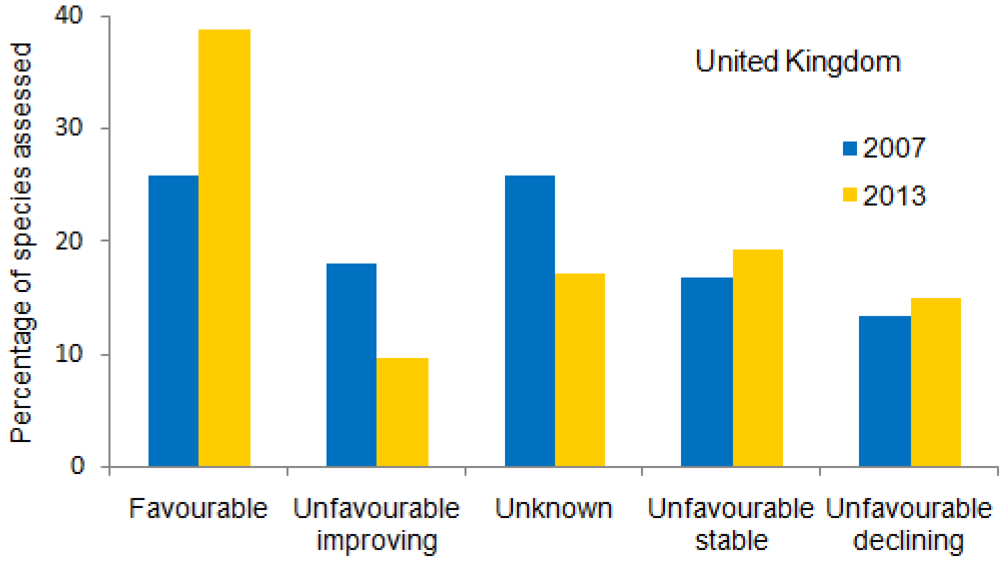


Fig.1. The C3b Status of UK species of European Importance indicator.

Appendix 2: General evaluation matrix for assessing conservation status of a species – taken from Annex C of the Explanatory Notes & Guidelines provided to EU Member States for the 2007-2012 Article 17 Reporting

Parameter	Conservation Status			
	Favourable ('green')	Unfavourable – Inadequate ('amber')	Unfavourable - Bad ('red')	Unknown (insufficient information to make an assessment)
Range	Stable (loss and expansion in balance) or increasing AND not smaller than the 'favourable reference range'	Any other combination	Large decline: Equivalent to a loss of more than 1% per year within period specified by MS OR more than 10% below favourable reference range	*No or insufficient reliable information available*
Population	Population(s) not lower than 'favourable reference population' AND reproduction, mortality and age structure not deviating from normal (if data available)	Any other combination	Large decline: Equivalent to a loss of more than 1% per year (indicative value MS may deviate from if duly justified) within period specified by MS AND below 'favourable reference population' OR More than 25% below favourable reference population OR Reproduction, mortality and age structure strongly deviating from normal (if data available)	*No or insufficient reliable information available*
Habitat for the species	Area of habitat is sufficiently large (and stable or increasing) AND habitat quality is suitable for the long term survival of the species	Any other combination	Area of habitat is clearly not sufficiently large to ensure the long term survival of the species OR Habitat quality is bad, clearly not allowing long term survival of the species	*No or insufficient reliable information available*
Future prospects (as regards to population, range and habitat availability)	Main pressures and threats to the species not significant; species will remain viable on the long-term	Any other combination	Severe influence of pressures and threats to the species; very bad prospects for its future, long-term viability at risk.	*No or insufficient reliable information available*
Overall assessment of CS	All 'green' OR three 'green' and one 'unknown'	One or more 'amber' but no 'red'	One or more 'red'	Two or more 'unknown' combined with green or all "unknown"

Fig.2. Evaluation matrix for assessing conservation status of a species.

Production of the index

Individual species accounts can be accessed via the JNCC's website (http://jncc.defra.gov.uk/page-6391) and the methods are discussed in the UK Approach document (http://jncc.defra.gov.uk/pdf/A17_2013_UKApproach.pdf).

For species, favourable conservation status is defined in Article 1(i) as when:

i. population dynamics data on the species concerned indicate that it is maintaining itself on a long-term basis as a viable component of its natural habitats, and;

ii. the natural range of the species is neither being reduced nor is likely to be reduced for the foreseeable future, and;

iii. there is, and will probably continue to be, a sufficiently large habitat to maintain its population on a long term basis.

To determine status using these considerations, it is necessary to know the current range together with some favourable reference range, the current population size together with a favourable reference population size, trends in the range and population size, and threats to the species and its habitat. The methods used to formalise this knowledge are necessarily varied, depending on the state of knowledge regarding the species and its requirements. The 2013 Approach document (see link above) sets out, in considerable detail, the considerations made in respect of the different types of knowledge available. The methodology is as clear as is possible, given the variety of circumstances to which it is applied. The categories of status are also well defined, with clear guidelines on when to attribute favourable, unfavourable or unknown status. The JNCC website notes that: "JNCC and the Country Conservation Bodies put a huge amount of effort into checking the assessments, and therefore have a high level of confidence that they are correct, and that changes, including within category changes, have been consistently and accurately discriminated."

In spite of this, the methodology is open to some inaccuracy and interpretation. This is partly because much of the data are not of high quality. For example, for terrestrial species, range often relies on data from the NBN, which is ad hoc with no reference to effort; for marine species, the data are often sparser. In addition, production of the index relies on a great deal of expert opinion. Some examples of where expert opinion is important include:

- When determining current range from 2007-2012, "for most species the date class was much wider, dictated by data availability and an expert understanding of current species distribution".
- Current range was often determined by fitting polygons around distribution data. How tightly the polygons were fitted to the data depended on a buffer parameter, at least partly determined by expert judgement.
- To determine population size, data were not always available from the 2007-2012 period. Where older data were used, they were "considered to be representative of the current population".
- Some estimates of both range and population size (by Method 1) relied on "expert opinion with no or minimal sampling".
- Population trends were based on robust surveillance where possible but, otherwise, were based on expert opinion.

- Population trends were only reported where "believed to be genuine".
- Where population trends in the 4 countries of the UK differed, and depending on the distribution of the species among those countries, the overall UK population trend was judged on a case-by-case basis.
- Whether the available habitat can support a viable population is not obviously based on quantitative approaches in many cases, and so is presumably subject to expert judgement.

As noted above, the extent of expert judgement obviously depends on the state of knowledge about the focal species. The guidelines are as comprehensive and as rigorous as is possible, given the range of species, life histories, habitats, requirements and threats. Moreover, the individual reports are rigorously laid out to a standard protocol, and are clear about which methods have been used throughout.

Interpretation

The indicator has a reasonably simple interpretation: ideally, after each reporting interval, the proportion of species assessed and deemed to have 'Favourable' status will increase. This might be at the expense of proportions of species in any other category but, again ideally, the proportions of species in the 'Unknown' and 'Unfavourable declining' categories would both be reduced at each successive reporting interval. Of course, reductions in the proportion of species in the 'Unknown' category might lead to increases in the proportions in other undesirable categories – but this must be viewed as an improvement – reflecting, as it does, an improvement in baseline data and, thus, an improvement in the state of knowledge about the species that triggered the change.

This is a primary indicator for the Aichi strategic goal C ("To improve the status of biodiversity by safeguarding ecosystems, species and genetic diversity"). It is specifically used to assess progress towards Target 12 ("By 2020, the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained"). As such, and because it is one of the few primary indicators that maps directly to an Aichi Target, this is a very important indicator.

Conclusions

It appears that the underlying production of the index is as robust as data permit. However, three developments might be worth considering.

Improvements to data collection

It is obviously important to strive to improve the quality of data underlying each species' assessment, especially those that are currently based on Method 1 ("estimate based on expert opinion with no or minimal sampling") for range size, population size, or both.

Improvements to analysis methods

It might be advantageous to implement a formal, quantitative framework for the assessment of viability, if and where this is currently lacking. There is an extensive literature on population viability and, whilst the current approach seems sensible given constraints on data for many species, it also seems likely that estimates of favourable reference ranges and population sizes could be determined with greater confidence (or, at least, that quantitative evidence could be supplied to justify the adequacy of existing 'favourable' reference points).

Improvements to interpretation

Although uncertainty is acknowledged wherever relevant in the individual species' assessments, each species is ultimately reduced to a number in a single category (see Fig. 1), regardless of the extent of uncertainty about that designation. i.e., uncertainty is accounted for in the underlying assessments but does not propagate through to the index itself. It might be possible to allocate species to more than one category, with weightings based on the probability with which the species falls into that category. How those weightings could be determined is a subject for discussion with the JNCC. However, techniques such as fuzzy logic (e.g. Cheung et al. 2005) or multiple 'blind' assessments (i.e., repeating the process with independent assessors to determine the robustness of the final designation) might present opportunities for this.

References

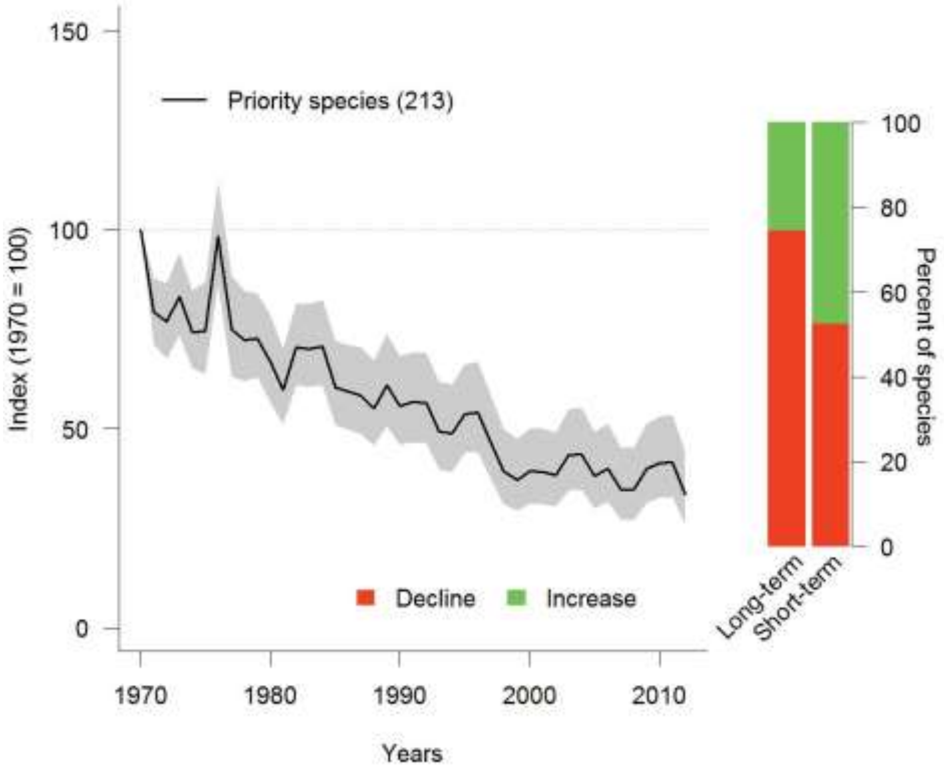
Cheung, W., Pitcher, T., Pauly, D. (2005) A fuzzy logic expert system to estimate intrinsic extinction vulnerabilities of marine fishes to fishing. Biological Conservation 124: 97-111

C4a - Status of priority species – relative abundance

Stephen Buckland

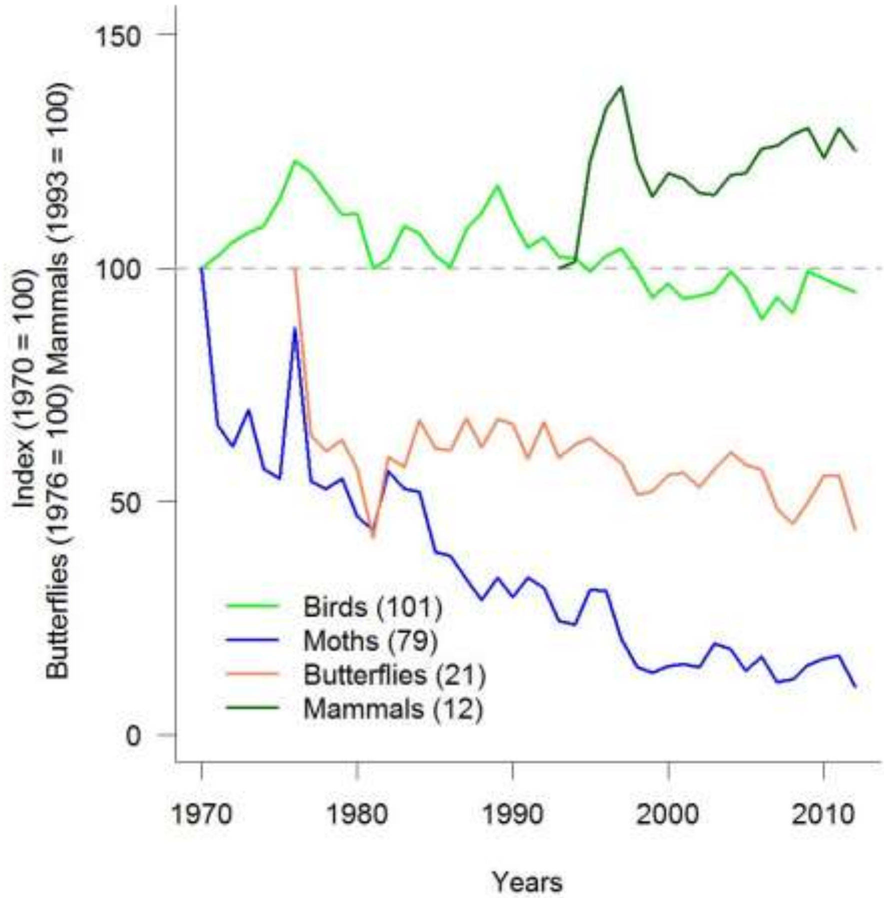
Background

This indicator is intended to represent trends in 2890 species, but only 213 species contribute to the index, and these are selected according to availability of suitable data. These 213 species include no species of plant, fish or marine mammals; by contrast, 101 species of bird (out of a possible 127), 79 species of moth, 21 species of butterfly and 12 species of terrestrial mammal are included. The composite index (a geometric mean of relative abundance, with 1970 as the baseline year, with relative abundance of 1), shows a value of 0.33 for 2012. This is a 67% decline in 42 years, although it has been relatively stable since the late 1990s.



The index for birds is largely stable over the time period; decreases in some communities (such as farmland) are offset by increases in others (such as waterbirds). The index for mammals (mostly bats) only starts in 1993. It shows a 38% increase between 1994 and 1997, but is otherwise fairly stable. The index for butterflies only starts in 1976. It shows a 40% reduction in the first year. Subsequently, it is largely stable, with a slight decline in recent years. The index for moths decreased by about 40% from 1970 to 1971. Fairly steep decline continued to 1998, by which time the index had decreased by around 86%. Subsequently, the index has been largely stable.

The decline in the overall index is driven primarily by the decline in moths. They account for 79 of the 213 species included, and many of those 79 species show big declines.



Data Quality

Survey design

This indicator combines data from many surveys. Even so, the indicator is intended to represent trends in 2890 species, but data from only 213 species contribute to the index. Of the 1787 plant species on the list, none are included in the indicator. Similarly, of 57 fish species, none are represented, and of 22 marine mammal species, none are represented. By contrast, of 127 bird species, 101 are represented, with data on any given species coming from at least one of six surveys.

Of the different survey schemes, survey design ranges from good (e.g. Breeding Bird Survey, stratified random sample) to subjective/purposive (e.g. moth survey, BMS, hedgehog road survey).

Where are the surveys done?

The Rothamsted moth survey is based on subjectively chosen sites. If you choose the best sites to monitor, you can expect deterioration over time (an example of regression to the mean). The surveys are based on light-traps, which attract an unknown proportion of moths on the ‘plot’, and the proportion varies by species. Plots are selected subjectively and opportunistically.

The Butterfly Monitoring Scheme is also based on non-random sites, and the transects within the sites sample the best habitat. All butterflies entering a ‘box’ ahead of the observer are in principle counted. In reality, some butterflies may remain undetected in the box, especially in cooler conditions. The greater activity of individuals in hot weather tends to bias counts upwards.

The dormouse survey is conducted at known sites. Nest-box counts are restricted of course to those animals using nest-boxes.

For bats, roost counts are said to be random, but are not in reality. Bat pass surveys are conducted along transects which tend to be located in edge habitats where bats are most numerous.

Hedgehog road surveys are non-random. Footprint tunnel surveys may allow some form of random sampling in the future.

Bird surveys are covered elsewhere in this review.

Changes over time in survey plot location

For moths plots tend to stay the same with little turnover (at least in the last 30 years), and span the country. For dormice, there is also little turnover of sites. Presumably there is higher turnover of sites and participants for hedgehog road surveys. Bats, birds and butterflies are covered elsewhere in this review.

Fieldwork methods

Moth light-trap surveys by their nature are restricted to night-flying species attracted to light. Equipment is standardised, and methods are straightforward. Dormice surveys are nest-box counts, which may fail to reflect any expansion of range. Hedgehog surveys involve driving a section of road, recording road kills.

Data quality assurance

There are multiple data sources. Bats, birds and butterflies are covered elsewhere. Moths are collected, so aside from occasional mis-identifications, the counts are presumably reasonably accurate. Presumably there is not a large data quality issue for dormice or dead hedgehogs.

Rigour of Analytical Method

A confidence interval for the indicator is calculated by bootstrapping, in which time series on individual species are resampled. No allowance for unrepresentativeness of sampling locations is made. The indicator is a composite of trends from a number of surveys, with varying statistical rigour.

Precision and Bias

The decline in moth species is believed to be real. Analyses to assess whether fewer moths might be attracted to traps due to greater light pollution suggest that this is not the case, as sites without an increase in light pollution show similar decline to sites with an increase (Conrad *et al.* 2004).

If all macro moths are considered, the decline in counts in the Rothamsted moth survey from 1968 to 2002 is 32% (see http://www.rutlandwater.org.uk/the-rothamsted-light-trap-network/). This decline is still substantial, but much smaller than the 86% decline in the index. This raises questions about the selection of species for the index. For example, if they were selected on the basis of their abundance in 1970, and if the moth community is naturally volatile, or responds rapidly to environmental change, then this is another example of regression to the mean. There should be some process for revising the list of species included, to avoid over-sampling declining species and under-sampling increasing species that were too rare at the outset to be included. Indeed, there might be a case for using the trend in total count of macro moths (corrected for number of traps) in the index, rather than the species-specific trends.

Moths account for 79 of the 213 species included in the index. Is it appropriate to give moths a weight of 37%? It would no doubt be difficult to agree other weightings, but I am not convinced that an index that is dominated by moths is useful. This strategy seems especially dubious, given that the moth species included are on average declining much more rapidly than moths in general. Birds are over-represented to an even greater degree than are moths (101 of the 213 species that contribute to the index), but their impact on the overall index is not as dramatic, because there does not appear to be the same bias towards declining species that the moths index exhibits. As 72 of the 79 moth species are also used in the frequency of occurrence version of the index, this apparent bias affects both indicators.

Large declines in moths have been observed in the south, whereas the average trend of monitored species is stable in Scotland. This may reflect climate change. If previously-rare species are increasing in the south, and these are not included in the 79 species analysed due to lack of records early in the time series, then this might explain in part the observed trends. Species previously rare in Scotland might have been common in the south at the outset, and hence included among the 79 species, thus removing the bias due to only monitoring the species that were initially common from the index for Scotland.

Because BMS is based on non-random sites, and the transects within the sites sample the best habitat, there is the potential for bias in either direction. Trends might be more favourable in the best sites, many of which may be protected, than in the wider countryside. However, by selecting the best sites to monitor, and then the best habitat within them, over time, the quality of these sites, and of the habitat along the transects, may decline (another example of regression to the mean). This might result in downward bias in trend estimates. BMS shows a substantial decline from 1976, the first year in which butterflies were included, to 1977. This is attributed to environmental conditions, but it is far from clear whether butterflies were abnormally abundant in 1976, or whether their abundance that year was representative of earlier years. Also, statistical artefacts can occur at the start of a new survey, when all observers are new.

The dormouse survey is conducted at known sites. If there is turnover in sites, such a strategy underestimates trends, so that the survey shows decline even if numbers are stable.

For bats, roost counts are claimed to be ‘random’. When all roosts are known, bias is unlikely. However, if there is turnover of colonies, and many roosts are unknown, then, as for dormice, trends will be underestimated. This is because a proportion of monitored colonies will disappear, and there is no mechanism for incorporating the increase from zero experienced by new colonies formed by movement from old colonies. Surveys based on counting bat passes using detectors are not subject to this bias, and as expected, they show more optimistic trends. There may be an issue of improving technology in the early days of bat detectors, which might result in upward bias in trends in bat pass counts. Perhaps this explains the steep rise in the mammal index (which is dominated by bats) in the first few years that mammals were included.

Numbers of hedgehogs found dead on roads are affected by traffic trends. An increase in road kill might reflect an increase in traffic, rather than an increase in abundance. It seems inappropriate to include such data.

Generally, biases would be expected to be small for most of the bird species included. However, the Wetland Bird Survey (WeBS) may produce biased trends for species whose pattern of usage of monitored sites may have changed, such as cormorants. Most surveys of rare birds are based on known sites, which will result in downward bias in trends for species that experience turnover in sites.

The inclusion of very rare species is in any case problematic. If there are any zeros in a time series, the geometric mean cannot be evaluated. The solution adopted for rare birds was to add 1% of the average value for a given species to all values in the time series. Further, if the index increased above 100 or fell below 0.01, it was set to 100 or 0.01 respectively.

Interpretation

Bootstrap confidence intervals are calculated to assess change over the long term (1970-2012), medium term (2002-2012) and short term (2007-2012). If the confidence interval includes zero, the indicator is assessed as 'no change'. Decrease was found in the long term and medium term, and no evidence of change in the short term ('no change'). It is noted that moths are influential; if they are omitted, the conclusion is 'no change' for all three time periods. The desirability of smoothing the species-specific time series prior to calculating the geometric mean is noted, and it is hoped that this can be implemented in future.

Conclusions

We conclude that the quality, clarity and relevance of this indicator could be improved in the following ways:

Improvements to data collection

We believe that this indicator has little value unless it becomes possible to include a more representative sample of species. This would involve a considerable additional sampling effort. Perhaps including a small number of species from each of the main taxa would be feasible, and might yield lower bias (but worse precision).

Improvements to analysis methods

We do not believe that changes in analysis methods can resolve the problems with this indicator. However, weighting the index to reduce the influence of moths and birds might give a better guide to trends in priority species. Weights can be determined from the percentage of priority species in each taxon that is included in the index. Those taxa with no available time series remain an issue, as does the apparently unrepresentative set of moth species. The criteria for selection of moth species might be reviewed.

Improvements to interpretation

The species-specific trends are combined using a geometric mean of relative abundances. It is sometimes interpreted as if it measures changes in average abundance. However, because it averages (on a log scale) relative abundances, even if total number of individuals is constant, the indicator can show a significant trend – downwards if common species tend to be increasing while rare species tend to be decreasing, and upwards if the converse is true. The indicator thus reflects trends in both abundance and evenness (Buckland et al. 2011). Use of additional subsidiary indicators to allow these two components to be separated should be considered.

The indicator should be interpreted in the context that different taxa represented on the priority species list show markedly different trends, and sampling intensity various hugely across taxa, from 0% (fish, plants, marine mammals) to 80% (birds).

References

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.

Conrad, K. F., Woiwod, I. P., Parsons, M., Fox, R. & Warren, M. S.**,** 2004. Long-term population trends in widespread British moths. *Journal of Insect Conservation* 8: 119–136.

C4b - Status of UK Priority Species – frequency of occurrence (insects)

Helen Roy

Background

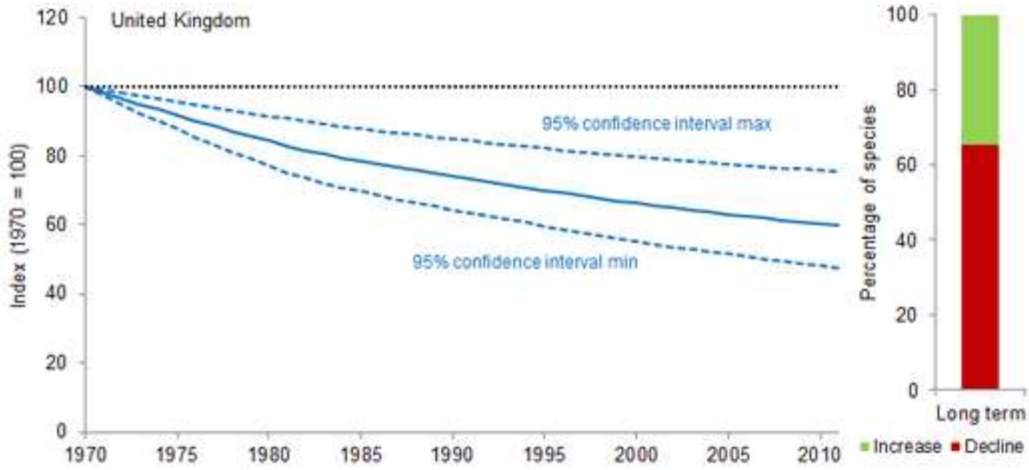
There is no longer reporting at the UK level on the status of species previously listed by the three-yearly UK Biodiversity Action Plan (UK BAP). Indicator C4b. infers 'threatened species' by using 'species identified as conservation priorities' within various country lists as a proxy. An inclusive approach was adopted, whereby a species only has to be included in one of the country lists to be included on the combined list.

The indicator shows changes in the frequency of occurrence of priority species (on one or more lists of each UK country) in the UK. The combined list includes 2890 species and 624 of these are insects but there is only sufficiently robust quantitative time series information (frequency of occurrence) for 179 species: 110 moths, 37 bees, 23 wasps, 2 ants, 2 hoverflies, 2 dragonflies and 3 grasshoppers. For most of the taxonomic groups the subset represents more than half of the species on the full country list (Table 2) but for hoverflies and ants the representation is poor. This is a composite indicator drawing on various datasets from the four countries (Table 2 = Table 3 in C4b. Status of UK priority species: Status of priority species – frequency of occurrence – insects Technical background report).

Table 2. Summary of species time series included in the Priority Species Frequency of Occurrence Indicator (FCL = Full Country List). Taken from Technical background report – Table 3)

Group	Data Type	Species with data	Species on FCL	Species on FCL with data and meeting criteria
Moths	Frequency of Occurrence	743	174	110
Ants		30	10	2
Bees		198	60	37
Wasps		201	33	23
Hoverflies		209	29	2
Dragonflies		39	4	2
Grasshoppers		31	6	3
Total included in indicator: 179				

The indicator shows decline in the frequency of priority species from 1970-2011.



Data Quality

Survey design

Biological records provided by the national schemes and societies are used for this indicator. Species selection is described and pragmatic – largely constrained by available information. A range of methods are now available for deriving trends from unstructured data such as biological records. There are a number of papers that describe these methods (Roy et al., 2012; Isaac et al. 2014; Isaac and Pocock, in press). The moths are the only group for which abundance and frequency of species occurrence data is available. It should be noted that 72 moth species are included both within C4b and C4a

Where are the surveys done?

Biological records are collected across the UK but the intensity of recording varies in space and time. There are inherent biases within such datasets (Isaac and Powney, 2015). The technical notes acknowledge that there are species on the priority lists for which data are too sparse to robustly model trends.

Changes over time in survey locations

Turnover of locations is likely to vary over time and across the different schemes and societies, however the “well-sampled sites model” is employed (Roy et al., 2012). Site visit is defined as a unique combination of data and 1km^2 grid cell.

Fieldwork methods

Various techniques are used by the volunteer recorders to gather the 1 km^2 (grid cell) occurrence data.

Data quality assurance

Taxonomic experts (volunteer and professional) within the national schemes and societies provide quality assurance. Additionally filters are used so that all visits with list lengths shorter than the median for the taxonomic group in question are excluded. Grid cells that have visits in less than three years are excluded.

Rigour of Analytical Method

Time series for each species is estimated from a generalised linear mixed effects model (year as covariate and grid cell as random effect (Roy et al., 2012) but testing of Bayesian occupancy models indicates high potential as an alternative method. The “well-sampled sites model” indicator is the geometric mean of the annual fitted values taken from the species specific “well-sampled sites model” linear models. The Bayesian indicator is the geometric mean of the species-specific annual estimates in the proportion of occupied sites after accounting for variation in detectability. The Bayesian model is concluded to be more understandable – presenting the proportion of sites occupied by a species, whereas the “well-sampled sites model” is the probability of observing a species on an average visit. The Bayesian model will enable assessment of change over shorter time periods (perhaps by decades).It is acknowledged that further work is still required to assess whether short-term fluctuations are reality or artefacts.

Precision and Bias

It is important to note that four of the insect groups (ants, hoverflies, dragonflies, grasshoppers) are represented by low numbers of species but these represent similar proportions of the total number of species on the four countries lists as the other groups.

95% confidence intervals were calculated using bootstrapping; in each iteration (n=10 000) a random sample of species were selected with replication and the geometric mean calculated. Bootstrapping at the site level rather than the species level is considered more robust but the data used with C4b are not derived from repeat site visits. Potentially a post hoc stratification of squares (1km^2 grid cell) method could be employed to enable bootstrapping at the site level.

A key assumption of the “well-sampled sites model” is species detectability does not change over time – the indicator accounts for this by excluding species for which taxonomic experts consider this assumption unsupportable. Extremely large or small index values can disproportionally influence the composite indicator so methods from C5 (wild bird index) were adopted (for index values that dropped below one). It should be noted that equal weighting was used throughout. Changes in the trend estimate for individual species have large impacts on the overall index. The index is sensitive to small sample-size effects.

Interpretation

The indicator provides trends in frequency of occurrence (distribution) of priority insect species but this could be misinterpreted or used by some as a proxy for population trends. However, the relationship between distribution and population trends is not consistent for all species and is influenced by a number of demographic factors, particularly colonisation rates and habitat characteristics (Freckleton et al. 2005). In populations where colonisation rates are high there is a positive relationship between occupancy and abundance but at low colonisation rates there is no relationship. Additionally at low abundance, Allee effects and demographic stochasticity render relationships sensitive to local changes in density.

Positive abundance–occupancy relationships are reported as "*among the most general macroecological patterns*" (Webb et al. 2012). There is limited empirical evidence of a high prevalence of negative relationships in nature and the few examples seem to occur as a consequence of spatially aggregation of an abundant species (Webb et al. 2012). However, it is essential that relationships between occupancy and abundance are assessed to provide context for interpretation of indicators derived from frequency of occurrence. relationships sensitive to local changes in density. It is essential that relationships between occupancy and abundance are assessed to provide context for interpretation of indicators derived from frequency of occurrence.

It is also important to note the caveat in the accompanying Technical Background Report to C4b "*Regardless of advances in statistical techniques there are species on the priority species lists for which data are currently too sparse to model robust trends. This is for a variety of reasons, including rarity (few occupied sites), low detectability or few active recorders. In order for the indicator to be representative of all types of species on the biodiversity lists, a method of assessing the changing status of a sample of these remaining data-poor species will need to be considered*." Such methods will also have to recognise the challenges of inferring population trends from occurrence data for such spatially restricted species. However, the relationship between trends in abundance and range (frequency of occurrence calculated using occupancy modelling) for UK butterflies has been examined and shown to be strongly correlated (Fox et al. 2015).

The aim of the indicator is to fulfil obligations, as a result of the adoption of the Strategic Plan for Biodiversity (including the Aichi Targets) at the 10th Conference of Parties of the Convention on Biological Diversity, to report progress against Aichi Target 12:

Target 12: By 2020 the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained.

Conclusions

We conclude that the quality, clarity and relevance of this indicator could be improved in the following ways:

Improvements to data collection

Biological records provide the necessary taxonomic breadth for this indicator but structure could be achieved through post hoc stratification of squares.

Improvements to analysis methods

Bootstrapping at the site level rather than the species level could be considered by applying a post hoc stratification of squares (1km^2 grid cell) method could be employed to enable bootstrapping at the site level.

Bayesian occupancy models have potential as an alternative method and such an approach is recommended given the inherent biases in the occurrence data.

Improvements to interpretation

Assess the possibility of using available structured survey data to assess both changes in abundance and evenness in abundance trends across species to improve interpretation of the indicator.

References

Fox, R., Brereton, T.M., Asher, J., August, T.A., Botham, M.S., Bourn, N.A.D., Cruickshanks, K.L., Bulman, C.R., Ellis, S., Harrower, C.A., Middlebrook, I., Noble, D.G., Powney, G.D., Randle, Z., Warren, M.S. & Roy, D.B. (2015). The State of the UK's Butterflies 2015. Butterfly Conservation and the Centre for Ecology & Hydrology, Wareham, Dorset.

Freckleton, R. P., Gill, J. A., Noble, D., & Watkinson, A. R. (2005). Large-scale population dynamics, abundance–occupancy relationships and the scaling from local to regional population size. Journal of Animal Ecology, 74(2), 353-364.

Isaac, N.J.B., can Strien, A.J., August, T.A., de Zeeuw, M.P. & Roy, D.B. (2014) Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution* 5, 1052–1060.

Isaac N.J.B. and Pocock M.J.O. (2015) Bias and information in biological records. *Biological Journal of the Linnean Society* 115, 522–531.

Roy, H.E., Adriaens, T., Isaac, N.J.B., Kenis, M., Onkelinx, T., San Martin, G., Brown, P.M.J., Hautier, L., Poland, R.L., Roy, D.B., Comont, R., Eschen, R., Frost, R., Zindel, R., Van Vlaenderen, J., Nedvěd, O., Ravn, H.P., Grégoire, J-C., de Biseau, J-C. & Maes, D. (2012) Invasive alien predator causes rapid declines of native European ladybirds. Diversity and Distributions DOI: 10.1111/j.1472-4642.2012.00883.x

Webb, T. J., Freckleton, R. P. & Gaston, K. J. (2012). Characterizing abundance–occupancy relationships: there is no artefact. Global Ecology and Biogeography, 21: 952–957. doi: 10.1111/j.1466-8238.2011.00736.x

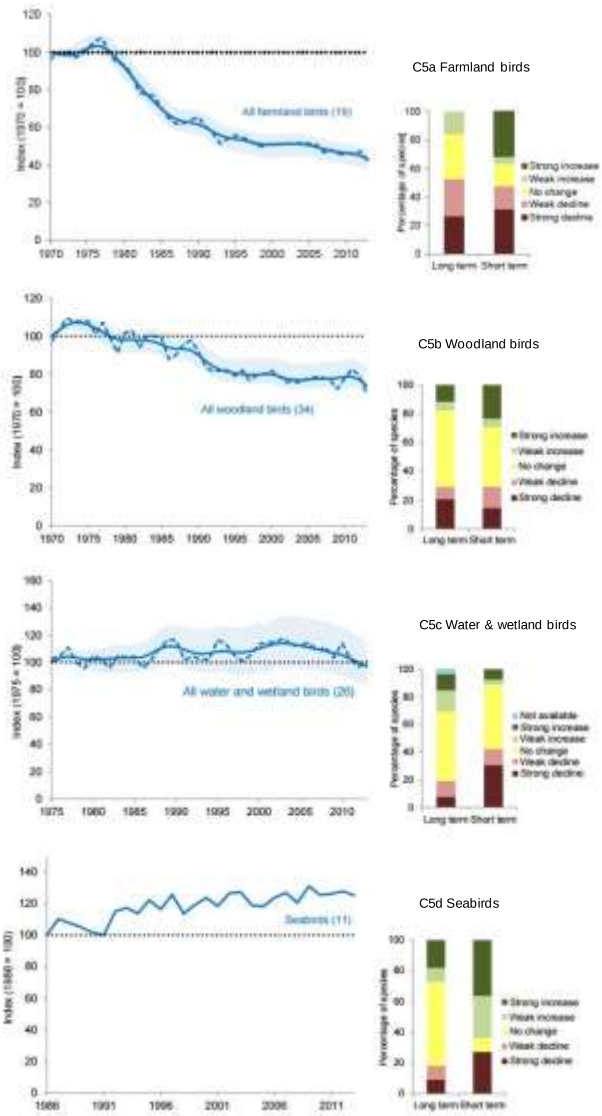
C5a-d - Birds of the wider countryside and at sea (excluding C5e wintering waterbirds)

Rhys Green

Background

This assessment covers C5a (farmland birds), C5b (woodland birds), C5c (wetland breeding birds) and C5d (seabirds). The indicators are generated by RSPB and BTO under contract to, and in collaboration with Defra and JNCC. The indicator includes sub-indicators of breeding bird populations in the four habitats listed above. A sub-indicator of wintering waterbirds (C5e) is assessed separately elsewhere. The four indicators are shown in Figure 1.

Figure 1. Indicators 5a-d for farmland birds, woodland birds, water and wetland breeding birds and seabirds. Graphs shows unsmoothed trend (dashed line) and smoothed trend (solid line) with its 95% confidence interval (light blue shading. Confidence intervals are not available for C5d. Bar chart shows the percentage of species within the indicator that have increased, decreased or shown no change, based on set thresholds of change.



C5a: farmland birds

Data Quality

Survey design

This sub-indicator is a multi-species indicator based upon breeding bird surveys of 19 species. Species were identified as predominantly associated with farmland from a previous classification based upon ecological information, checked by analysis of habitat-specific BBS survey data to allow the estimation of the proportion of the national population located in each habitat (Newson, Noble & Eaton 2004). The sub-indicator is based upon the same set of species in England and the UK. The sub-indicator is based upon combined analyses of two surveys, CBC and BBS, for all species except rook *Corvus frugilegus*. The data for rook are from both the BBS and two national rook censuses in 1975 and 1996. The BTO/JNCC Common Bird Census (CBC) was the first systematic survey of widespread breeding birds in the UK and ran from 1961 to 2000. CBC survey plots were selected by volunteer observers and were not representative of UK bird habitats in general. The field surveys were conducted by volunteer observers who followed a protocol designed by expert ornithologists. Raw data (maps of bird records within a survey plot from several visits) were processed and interpreted to estimate numbers of territories by professional staff using standard methods, though with some judgement. The BTO/JNCC Breeding Bird Survey (BBS) was designed to replace CBC and overlapped with it (1994 to present). BBS was designed by expert ornithologists and statisticians. It involves the observer walking, as nearly as possible, two parallel 1-km transects and recording all birds seen and heard. Distances of birds from the transect line are estimated. The survey is conducted twice in each square, in April and June. Results are processed by standard methods. The methods and survey site selection have been the subject of peer-reviewed papers in the scientific literature, though some details of the analytical methods are only available in the grey literature.

Where are the surveys done?

CBC and BBS are both surveys conducted on sample plots which together comprise a small subset of the total breeding habitat of the selected species within England and the UK. All plot locations are georeferenced and the data appropriately archived. CBC surveys were done on plots chosen and delineated by volunteer observers; often whole farms or woods. The plots were not a random or representative sample of locations in the UK. The mean area of farmland plots (about 70 ha) declined somewhat over time, but mostly in the early years of the programme. The main bias in farmland plot distribution was that there were disproportionately more plots in southern Britain. However, within southern Britain, farm types within plots were broadly similar to those elsewhere (Fuller, Marchant & Morgan 1985). Plot distribution was influenced by the distribution and preferences of volunteer surveyors. CBC plots were classified as farmland or woodland, based upon their predominant habitat, but the sub-indicator uses data from all plots for the set of species defined as farmland species. The choice of BBS survey plots (1-km squares) was random within strata (geographical regions), though the sampling fraction varies among regions so as to make use of variation in the availability of volunteer surveyors. The random sample was drawn from all 1-km OS grid squares within a region, excluding coastal squares with less than 50% land. Both CBC and BBS collected data on birds of all ages and both sexes. However, singing territorial males are probably predominant for many species. The CBC explicitly estimated the number of breeding territories, with song records playing an important part in estimating this. The national surveys of rooks were based upon sample surveys of occupied nests in randomly selected tetrads by BTO volunteers. Take up of rook survey squares was sufficiently high that the sample can be regarded as representative.

The methods and survey site selection for all of these surveys have been the subject of peer-reviewed papers in the scientific literature, though some details of the analytical methods are only available in the grey literature.

Changes over time in survey plot location

Both CBC plots changed over time as observers were recruited or left the scheme. Sometimes another observer took over a plot, but often not. Replacements were not necessarily similar to those lost. BBS plots may cease to be surveyed and new ones are introduced. However, efforts are made to maintain coverage of the same squares, even if the observer changes, and the selection process ensures that survey squares remain representative.

Fieldwork methods

CBC, BBS and rook survey field methods are written up as instruction protocols for volunteers and in peer-reviewed and grey literature reports. Surveyors are thoroughly instructed in techniques and can seek clarification from professional staff.

Surveyors are skilled in bird identification, though formal tests of this are not conducted. Most surveyors were recruited by BTO Regional Organisers, who are expert ornithologists. BBS surveyors are offered a CD of bird calls and given the opportunity to attend training courses and workshops. It is possible that age-related deterioration in the capacity of observers to hear high-frequency calls and songs might affect the proportion of birds detected and result in changes in detection over time. An analysis of detectability using BBS data on the distance from the transect at which detected birds were recorded found that detectability significantly declined over time since about 2000 for two of 20 common bird species examined (Newson et al. 2013). However, this is unlikely to be because of changes in hearing because neither of the species concerned have unusually quiet or high-frequency songs.

CBC and BBS surveys are conducted at times of year and times of day when detectability of most species is expected to be high. Multiple visits during the season are needed so that species with song peaks or migration arrival dates at different times are adequately sampled. The CBC, which had several visits spaced through the spring and summer was more robust in this respect than the BBS, which has two visits.

Data quality assurance

CBC, BBS and rook survey data are prepared for analysis using thorough and repeatable protocols. Efforts are made by professional staff to check for errors during collection and data entry. Programs are run to detect unusual records and these are checked.

Rigour of Analytical Method

Population trends were produced from CBC and BBS data using a log-linear Poisson regression model with count per plot as the dependent variable for each species. The analysis includes site and year effects. The site effect allows for turnover in sites and missing values. The back-transformed estimated year effects are annual indices of abundance for that species. Trends are generated from combined CBC and BBS data by maximising the joint likelihood. Trends based upon BBS are adjusted for regional differences in sampling fraction by weighting analysis by the ratio of the number of available squares in a region to the number covered. No such weighting is performed for the CBC data. CBC and BBS data are assigned equal weight in the analysis. Smoothing is done after annual values are generated for each species using a thin-plate smoothing spline procedure. Confidence limits on trends for individual species are generated by bootstrapping, with the survey plot being the bootstrap unit. For the BBS, the bootstrap unit is the 1-km square, not the transect sections within it. Geometric means of species' indices are used to generate the multi-species indicator. Confidence limits for the multi-species indicator are derived for the species-specific bootstrap results. BBS species trends are currently not adjusted for possible changes over time in detectability or for possible unrepresentativeness of habitat along transects, though detectability is taken into account when estimating national population sizes (Newson et al. 2008). Species are given equal weight in the calculation of the multi-species indicator from the smoothed species trends regardless of their abundance.

Precision and Bias

The sub-indicator is produced with appropriate bootstrap confidence intervals based upon bootstrapping of species' values by survey site. These show a reasonable level of precision for the detection of long-term (decadal) changes.

The selection of survey sites as a stratified random sample with adjustment of the results to allow for unequal sampling within strata is a robust approach to obtaining unbiased estimates. However, there are some potential sources of bias that are not currently adjusted for, though the importance of this cannot be fully assessed at present. BBS species trends are currently not adjusted for possible changes over time in detectability. An analysis of detectability using BBS data on the distance from the transect at which detected birds were recorded found that detectability declined significantly over time since about 2000 for two of 20 common birds examined (Newson et al. 2013). For these species, there was a substantial difference between the smoothed trends produced with and without adjustment for detectability changes. The two species concerned are among the most abundant examined and it is possible that the change was detected by significance testing for these and not others because the test had the greatest power for them. It was concluded that the findings of this study did not justify the introduction of routine adjustment for changes in detectability in the calculation of the indicator. However, this conclusion should be kept under review. Changes in detectability might also result from changes in the timing of arrival of breeding of birds, caused by climatic change. This might result in the timing of peak detectability and the timing of survey visits becoming increasingly mismatched. The BBS survey might be especially prone to this potential bias because it only has two survey visits. The detectability analyses of Newson et al. (2013) provided no indication of date-specific changes, as would have been expected from climatic change. However, detailed analyses were only done for two non-migratory species.

BBS transects in many survey squares do not follow the idealised transect routes because observers are reluctant or unable to do so because of physical barriers, ease of walking, disapproval of landowners and privacy considerations. Population changes might be larger in some habitats than others, so this discrepancy could bias trend estimates even if the amount of habitat discrepancy between ideal and actual transects stayed the same over time. Trends could be modelled with habitat characteristics of 200 m transect sections as covariates using, for example, the Freeman-Newson approach. This would allow the trends to be adjusted to allow for differences in habitat between the actual and ideal transect routes. This is not done because the required analyses would be complex and costly. It seems unlikely that such adjustments would make much difference, but no detailed evaluation of this potential bias has been performed. The topic should be kept under review.

Interpretation

Species-specific trends are combined using a geometric mean of relative abundances. The farmland bird indicator forms part of the evidence base which is synthesised for reporting, to inform a wider policy evaluation of progress towards the Aichi targets. It addresses, in part, Aichi Target 12 ("By 2020, the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained"). It is of direct relevance to Strategic Goals A, B and C of the UK Government's Post-2010 Biodiversity Framework.

Conclusions

We conclude that the quality of this indicator could be improved in the following ways:

Improvements to data collection

The current design of the surveys, which involves stratified random sampling, is satisfactory in allowing coverage of a representative range of sites. Survey routes within survey squares are not in most cases representative of the habitats within the squares as a whole and could not easily be made so. This is especially the case for farmland, where crops and boundaries make walking representative transects impractical. The habitat data collected on transect sections would allow adjustment for bias caused by this, if necessary. Recording of the distance of registration from the transect line allows for adjustment for detection probability, if necessary. Most of the more abundant farmland bird species of the UK are covered by the indicator. Rare farmland species are not included, though reliable data exist on some of these, including retrospectively. Consideration should be given to whether to include such species, though the choice of whether this is desirable or not is largely determined by a subjective view of what the indicator is intended to represent. It appears to be the case that rare farmland species are included in the indicator to a lesser extent than is the case for rare woodland birds.

Improvements to analysis methods

BBS species trends are currently not adjusted for possible changes over time in detectability, but the analyses described above suggest that this might bias trends to some extent and for some species and therefore bias the indicator as a whole. Bias potentially introduced by the effects of climatic change on bird phenology and detectability might occur and is not excluded by the analyses conducted so far. Existing analyses do not give rise to major concerns about bias in the indicator, but their coverage so far is limited. No changes to the analysis methods are justified, given current information, but the topic should be given further consideration.

BBS transects in many survey squares do not follow the idealised transect routes. Population changes might be larger in some habitats than others, so this discrepancy could bias trend estimates. The adjustment of trends to allow for differences in habitat between the actual and ideal transect routes is not done because the required analyses would be complex and costly. It seems unlikely that such adjustments would make much difference, but no detailed evaluation of this potential bias has been performed. The topic should be given further consideration.

Improvements to interpretation

See also the generic issues relating to the use of composite trends, and the difficulties of differentiating changes in abundance from changes in evenness.

C5b: woodland birds

Data Quality

Survey design

This sub-indicator is a multi-species indicator based upon breeding bird surveys of 38 species of woodland birds. Species were identified as predominantly associated with woodland from a previous classification based upon ecological information, checked by analysis of habitat-specific BBS survey data to allow the estimation of the proportion of the national population located in each habitat (Newson, Noble & Eaton 2004). Three of the species covered by the UK version of the indicator are absent (capercaillie *Tetrao urogallus*) or not sufficiently sampled (crossbill *Loxia curvirostra* and pied flycatcher *Ficedula hypoleuca*) in England so they are not included in the England version. The sub-indicator is based upon combined analyses of two surveys, CBC and BBS, for all species except capercaillie, for which data from national (Scottish) transect surveys are used.

The assessment of the CBC and BBS survey design is the same as for sub-indicator C5a (see above).

Where are the surveys done?

The assessment of the CBC and BBS plot location is similar to that for C5a. The CBC woodland plots were not a random or representative sample of locations in the UK. The main bias in plot distribution was that there were disproportionately more plots in southern Britain. CBC plots were classified as farmland or woodland, based upon their predominant habitat, but the sub-indicator uses data from all plots for the set of species defined as woodland species. The capercaillie survey data are from a series of surveys based upon a stratified random sample of line transects.

Changes over time in survey plot location, fieldwork methods and data quality assurance

The assessment of the CBC and BBS surveys for these questions is the same as for sub-indicator C5a (see above).

Rigour of Analytical Method and Precision and Bias

The assessment of the CBC and BBS surveys for these questions is the same as for sub-indicator C5a (see above).

Interpretation

Species-specific trends are combined using a geometric mean of relative abundances. The woodland bird indicator forms part of the evidence base which is synthesised for reporting, to inform a wider policy evaluation of progress towards the Aichi targets. It addresses, in part, Aichi Target 12 ("By 2020, the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained"). It is of direct relevance to Strategic Goals A, B and C of the UK Government's Post-2010 Biodiversity Framework.

Conclusions

We conclude that the quality of this indicator could be improved in the following ways:

Improvements to data collection

The current design of the surveys, which involves stratified random sampling, is satisfactory in allowing coverage of a representative range of sites. Survey routes within survey squares may not be completely representative of the habitats within the squares as a whole but this seems less likely to be a problem than for farmland. The habitat data collected on transect sections would allow adjustment for bias caused by this, if necessary. Recording of the distance of registration from the transect line allows for adjustment for detection probability, if necessary. Abundant and quite scarce woodland bird species of the UK are covered by the indicator. A few rare woodland species are not included, though reliable data exist on some of these, including retrospectively. Consideration should be given to whether to include such species, though the choice of whether this is desirable or not is largely determined by a subjective view of what the indicator is intended to represent.

Improvements to analysis methods

BBS species trends are currently not adjusted for possible changes over time in detectability, but the analyses described above suggest that this might bias trends to some extent and for some species and therefore bias the indicator as a whole. Bias potentially introduced by the effects of climatic change on bird phenology and detectability might occur and is not excluded by the analyses conducted so far. Existing analyses do not give rise to major concerns about bias in the indicator, but their coverage so far is limited. No changes to the analysis methods are justified, given current information, but the topic should be given further consideration.

BBS transects in squares with large blocks of woodland are more likely to be close to the idealised route than those in squares that are predominantly farmland. Hence, the potential for bias is likely to be smaller than for farmland. It therefore seems unlikely that such adjustments would make much difference.

Improvements to interpretation

See also the generic issues relating to the use of composite trends, and the difficulties of differentiating changes in abundance from changes in evenness.

C5c: wetland breeding birds

Data Quality

Survey design

This sub-indicator is a multi-species indicator based upon surveys of 26 bird species characteristic of freshwater wetlands, including waterways, reedbeds and wet meadows. Species were identified as predominantly associated with wetlands from a previous classification based upon ecological information. The indicator runs from 1975 onwards. Nine of the species were added to the set after the inception of the indicator at various times between 1977 and 2004. The indicator is calculated using more datasets than C5a and C5b and differs from them in that it uses survey data specific to wetland habitats, combined with data from all habitats from CBC and BBS, whereas C5a and C5b use data from all habitats for species designated as characteristic of farmland and woodland respectively. The surveys used comprise the Waterways Bird Survey (WBS), Waterways Breeding Bird Survey (WBBS), CBC and BBS (see above), constant-effort mist-netting at wetland sites (part of the Constant Effort Scheme, CES) and special surveys of the little egret *Egretta garzetta*.

Where are the surveys done?

The diverse array of surveys used in the calculation of this sub-indicator make a concise assessment of survey plot selection difficult. WBS and CES sites are selected by volunteers and may not be representative of all wetland breeding habitats of the species concerned. WBBS and BBS plots are selected using a stratified random sampling approach similar to that described for C5a and C5b.

Changes over time in survey plot location, fieldwork methods and data quality assurance

As for CBC and BBS surveys, there is turnover in plot locations for WBS, WBBS and CES.

Rigour of Analytical Method and Precision and Bias

The assessment of the CBC and BBS surveys for these questions is similar to those for sub-indicator C5a and C5b (see above). However, it is not clear from available documents how weighting was done to allow for variation in sampling fraction among geographical regions (strata).

Interpretation

Species-specific trends are combined using a geometric mean of relative abundances. The wetland bird indicator forms part of the evidence base which is synthesised for reporting, to inform a wider policy evaluation of progress towards the Aichi targets. It addresses, in part, Aichi Target 12 (“By 2020, the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained”). It is of direct relevance to Strategic Goals A, B and C of the UK Government’s Post-2010 Biodiversity Framework.

Conclusions

We conclude that the quality of this indicator could be improved in the following ways:

Improvements to data collection

The heterogeneous nature of the surveys contributing to this indicator makes ensuring representativeness of data collection difficult. Periodic breeding season surveys of a stratified random sample of squares containing wet features, resembling the BTO’s Dispersed Waterbirds Survey of wintering wetland birds, might be considered to remedy this and would allow coverage of a larger representative range of sites than the BBS allows. Abundant and quite scarce wetland bird species of the UK are covered by the indicator. A few rare wetland species are not included, though it is doubtful whether reliable data exist on these.

Improvements to analysis methods

Further assessment of how weighting is done to allow for variation in sampling fraction among geographical regions (strata) would be useful, as would more detailed consideration of how the results from the different surveys are combined.

Improvements to interpretation

See also the generic issues relating to the use of composite trends, and the difficulties of differentiating changes in abundance from changes in evenness.

C5d: seabirds

Data Quality

Survey design

Counts of UK seabirds at breeding colonies are collected annually by the Seabird Monitoring Programme (SMP). In addition, there were attempts to conduct complete censuses of all breeding seabirds in the UK in 1985-1988 and 1998-2002. This sub-indicator is a multi-species indicator based upon surveys of breeding bird surveys of 14 species at the UK level in 2014, reduced, perhaps temporarily, to 13 species in 2015. The 14 species included in 2014 were Northern fulmar *Fulmarus glacialis*, European shag *Phalacrocorax aristotelis*, great cormorant *Phalacrocorax carbo*, Arctic skua *Stercorarius parasiticus*, Arctic tern *Sterna paradisaea*, black-legged kittiwake *Rissa tridactyla*, common guillemot *Uria aalge*, common tern *Sterna hirundo*, great black-backed gull *Larus marinus*, herring gull *Larus argentatus*, lesser black-backed gull *Larus fuscus*, little tern *Sternula albifrons*, razorbill *Alca torda*, and Sandwich tern *Sterna sandvicensis*. The indicator previously covered a wider range of species but was modified to include fewer in 2014. Exclusion of seabird species from the indicator is principally because of (a) insufficiently precise annual counts, (b) counts which cover an inadequate fraction of the population, (c) mismatch between trends determined by the imputing method on the sample of colonies and trends from complete censuses and (d) insufficiently reliable counting methods. The UK has 23 seabird species with significant breeding populations. The seabird species not included in the indicator are Manx shearwater *Puffinus puffinus*, northern gannet *Morus bassanus*, storm petrel *Hydrobates pelagicus*, Leach's petrel *Oceanodroma leucorhoa*, great skua *Stercorarius skua*, roseate tern *Sterna dougalli*, Atlantic puffin *Fratercula arctica,* black guillemot *Cepphus grylle*, eider *Somateria mollissima*. For the purposes of this review, two possible species are not considered to be seabirds because they have significant inland populations (black-headed and common gulls *Larus ridibundus* and *L. canus*). Of seven seabird species for which the UK hold more than 10% of the global breeding population, four are included in the indicator, but three are not (Manx shearwater, northern gannet and great skua). Hence, the species coverage of this survey is notably incomplete.

Where are the surveys done?

Where possible, whole colonies or representative sample plots within colonies are counted annually. The counts cover a sample of all colonies. However, it is often not possible to perform counts annually and making sample areas representative is difficult. All survey site locations are georeferenced and the data appropriately archived. Many breeding colonies are difficult to access or view and some are remote. Hence, much of the range of sites covered by the survey for most species is probably determined more by opportunity and constraint than design.

Some details of these surveys have been the subject of peer-reviewed papers in the scientific literature. Some further details of the analytical methods are available in the grey literature. Coverage of as many as possible of the breeding sites thought to exist is attempted at intervals of about 20 years. Two of these surveys have been completed and a third is due for completion within the next few years.

Changes over time in survey plot location

There are many gaps in survey coverage of seabird colonies for annual or periodic monitoring and coverage has changed over time.

Fieldwork methods

Methods for counting seabirds vary widely among species and some species, especially the burrow and cavity-nesting species (e.g. petrels and shearwaters) are difficult to survey. Reliable methods for these species are still under development. Some species (e.g. terns) have breeding colony locations that shift annually or periodically, making it difficult or inappropriate to sample the same set of colonies repeatedly. In some cases, colonies may remain in place but large transfers of birds occur from one to another. For these species, an attempt must be made to cover as large a proportion of colonies as possible. The most tractable species for survey are those that nest in the open at fixed locations on cliffs or islands (e.g. guillemot, northern gannet). Photographic methods are feasible for counting open-nesting species.

Data quality assurance

Survey results are checked for obvious errors before inputting to databases.

Rigour of Analytical Method

Population trends with annual values were produced for each species from SMP population data. For colonies with missing data for a given year, an imputing method (Thomas 1993) was used to calculate missing values which were then included in the calculation of an annual index for the species values across the set of colonies. This method was tested for the heronry surveys for which it was developed, but is unclear whether the method and the choices of smoothing and weighting parameters used are appropriate for seabirds. It is not clear whether alternative methods, such as the sites x years Poisson regression analysis used for BBS data, would be similar or superior. Confidence limits of annual indices for each species were obtained by bootstrapping, with survey site being the bootstrap unit. Given that there are presumed complete survey data for all sites for two periods, one at the beginning of the time series (Seabird Colony Register Census) and another about half-way through it (Seabird 2000), it is not clear how the errors attached to those near-complete censuses (especially the second) are treated in the population index calculations. In theory, if a species had two near-complete surveys in (about) 1986 and (about) 2000 the confidence interval for the difference in population between these two years should be small, but the SMP data show that the intervals for the 1986-2000 change, representing the Seabird 2000 period, are unexpectedly large for most species. This requires further consideration. The species indices and multi-species index are not smoothed. The multi-species indicator does not have confidence intervals.

Precision and Bias

For the 19 species (among them some not included in the calculation of the UK indicator) with species-level UK population indices included in the SMP database, the ratio of the upper to the lower bound of the 95% confidence interval for the index of population size in 2013 relative to that in 1985, ranged from 1.2 to 16.0 (geometric mean ratio 2.8). For the 14 species included in the UK indicator, the range of these ratios 1.2 to 8.5 and the geometric mean ratio 2.5. Only for five of these species was the ratio less than two. The multi-species sub-indicator is produced without confidence intervals. Hence, the precision of species trends over long periods is poor and there is no formal estimate of precision for the multi-species indicator.

There are some potential sources of bias that are not currently adjusted for. Attendance at breeding colonies by adult seabirds is known to vary with environmental conditions and changes in this may result in bias. Incomplete and unrepresentative sampling of colonies and survey sites within colonies may introduce substantial bias. For species with mobile colonies, such as terns, it is difficult to keep track of the all the colonies, so apparent declines and increases could be spurious and result from loss or discovery by observers of colonies.

Interpretation

Species-specific trends are combined using a geometric mean of relative abundances. Because many UK seabird species are not included in the multi-species indicator, including some numerically abundant species, it provides a potentially unreliable indicator of seabird populations as a whole. Because there is considerable imputing of missing annual values, short-term changes in trends may be particularly unreliable.

The seabird indicator forms part of the evidence base which is synthesised for reporting, to inform a wider policy evaluation of progress towards the Aichi targets. It addresses, in part, Aichi Target 12 ("By 2020, the extinction of known threatened species has been prevented and their conservation status, particularly of those most in decline, has been improved and sustained"). It is of direct relevance to Strategic Goals A, B and C of the UK Government's Post-2010 Biodiversity Framework.

Conclusions

We conclude that the quality of this indicator could be improved in the following ways:

Improvements to data collection

It would be valuable to conduct a thorough review of the number and selection of breeding colonies to be counted and the frequency (annual or periodic) of survey. This should be done for all UK seabird species, including those not currently included in the indicator. This might allow the degree of bias caused by surveying of colonies with atypical trends to be reduced. A stratified random sampling procedure is difficult to design and implement because of the inaccessibility of some colonies, but should be attempted along with simulations of likely changes to bias and precision. Because many seabird species are long-lived and faithful to their breeding colonies, short-term fluctuations in breeding populations are rare in many cases. Hence, it may be possible to reduce survey frequency as a means of increasing the representativeness of colony coverage without a large increase in the resources required. New survey methods for some species might be used, including increased use of photographic and acoustic surveys (e.g. Oppell et al. 2014). The aim should be to reduce bias and improve precision of the compentent population indices for species already included in the indicator and also expand the number of species covered by it. It is highly unsatisfactory that the indicator of UK seabird populations does not include some of the species for which the UK holds a significant fraction of the global population.

Improvements to analysis methods

A review should be conducted of the appropriateness of the imputing method of Thomas (1993) for estimating missing annual counts so that a species-specific index can be calculated. Alternative methods, such as the sites x years Poisson regression analysis (as used for BBS data) and GAMs should be trialled. There are presumed near-complete survey data for all sites for two periods, one at the beginning of the time series (Seabird Colony Register Census) and another about half-way through it (Seabird 2000). Production of a new near-complete survey of UK seabird populations is in progress and should be expedited. The incorporation of data from these near-complete surveys into the production of the annual indicator requires careful re-appraisal. A method should be found to produce bootstrapped confidence intervals for the multi-species indicator.

Improvements to interpretation

Interpretation of this indicator is rendered problematic because it does not include several important species (see above) and the precision and bias of the indicator are unclear. This will only be remedied by steps to resolve the problems with data collection and analysis decribed above. See also the generic issues relating to the use of composite trends, and the difficulties of differentiating changes in abundance from changes in evenness.

References

Fuller, R.J., Marchant, J.H. & Morgan, R.A. (1985) How representative of agricultural practice in Britain are Common Birds Census farmland plots? Bird Study 32, 56-70.

Newson, S.E., Evans, K.L., Noble, D.G., Greenwood, J.J. & Gaston, K.J. (2008) Use of distance sampling to improve estimates of national population sizes for common and widespread breeding birds in the UK. Journal of Ecology, 45, 1330–1338.

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

Newson, S.E., Noble, D.G. & Eaton, M.A. (2004) Preliminary BBS-based habitat-specific indicators for wild bird populations: 1994-2002. BTO Research Report No. 368

Oppel, S., Hervías, S., Oliveira, N., Pipa, T., Silva, C., Geraldes, P., Goh, M., Immler, E. & McKown, M. (2014) Estimating population size of a nocturnal burrow-nesting seabird using acoustic monitoring and habitat mapping. Nature Conservation 7: 1-13

Thomas, G.E. (1993) Estimating annual total heron population counts. Appl. Statistics, 42, 473-486.

C5e - Wintering waterbirds

Stephen Buckland

Background

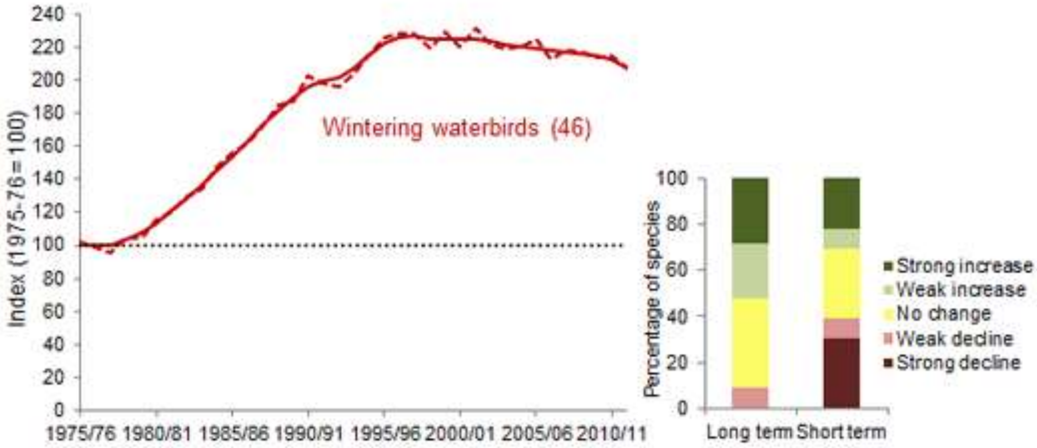
Wintering waterbirds comprise waders, ducks, geese, swans and other waterbirds such as cormorants and coot. Trends for waders and most wildfowl are derived from the Wetland Bird Survey (WeBS), although trends for several wintering geese populations (e.g. pinkfeet, Greenland whitefront and Svalbard barnacle geese) are derived from the Goose and Swan Monitoring Programme. Counts for WeBS are made at all wetland habitats, including freshwater lakes, ponds, reservoirs, gravel pits, rivers, canals and marshes as well as open coasts and estuaries. Counts are made once a month ideally on predetermined priority dates.

WeBS core counts provide the information used in assessing population trends. Annual summaries of the monthly counts are analysed each year using the Underhill Index method (Underhill and Prys-Jones 1994) specifically developed for waterbird populations, to produce a time series of index values for each species or subpopulation. This method includes a calculation to estimate counts for missing site-month combinations, based on counts in other months and all sites.

Data for wildfowl are available for the period 1966/67 to present. Data for waders are available only from 1974/75 onwards because a high proportion of counts before this winter were imputed. For species added later to the scheme (great crested grebe and coot in 1982/83, little grebe in 1985/86, cormorant in 1986/87), data from the first two years following their inclusion were omitted from indices as initial take-up by counters was incomplete. The UK wintering waterbird indicator incorporates population trends for 46 species, races or populations treated as separate components.

The index is formed by taking the geometric mean of species relative abundance trends. Those trends might be smoothed (using generalized additive models) or not, so that a smoothed and an unsmoothed index is obtained.

The index shows large increases from the mid-1970s to the mid-1990s, followed by a steady but small decline.



Data Quality

Survey design

The concept is to count the whole population, so there is not a design as such.

Where are the surveys done?

In principle, all wetland sites are counted. In practice, many minor sites are not counted. For some species, the counts are probably nearly complete counts of the wintering population. For other species, this is clearly not the case.

Changes over time in survey plot location

An attempt is made to count all major sites on every occasion.

Fieldwork methods

The method is to count all individuals at each site. Training workshops are available, but not required. At the major sites, counts tend to be carefully coordinated.

Data quality assurance

Most data are now submitted online, with standard checks.

Rigour of Analytical Method

Imputation of missing counts is done using the method of Underhll and Prys-Jones (1994).

Precision and Bias

For some species, the data can be expected to be fairly complete, with a high proportion of the population counted on each occasion. For other species, a relatively small proportion would be counted, and trends may not reflect trends in the total population. For example little grebe counts in cold winters, when much of their normal habitat is frozen, may be very different from those in mild winters. A trend in winter temperature might then bias trends in counts. Similarly, cormorant trends may reflect the increasing trend for cormorants to occur inland, more than the trend in the coastal population.

Many sites hold large numbers of birds which may be very difficult to count. Some bias can be expected which is likely to vary by site and by observer; this is probably not a large concern, unless there is a time trend in the bias.

Interpretation

The species-specific trends are combined using a geometric mean of relative abundances. The BTO website gives an excellent summary of potential interpretation pitfalls:

http://www.bto.org/volunteer-surveys/webs/publications/webs-annual-report/numbers-trends/methods/interpretation-waterbird-counts

However, this focuses on interpretation of species-specific trends, rather than on an index averaged across species.

Conclusions

We conclude that the quality, clarity and relevance of this indicator could be improved in the following ways:

Improvements to data collection

Greater use of replicate counts by observers operating independently might allow the reliability of the counts to be quantified.

Opportunistic independent verification of counts at some sites using a different method, for example high-resolution aerial imagery, possibly using drones, would be invaluable for quantifying bias.

Improvements to analysis methods

The imputation method used seems appropriate. However, some of the target species may not be adequately surveyed to allow reliable inference. We recommend that an assessment is carried out for each target species, to determine whether its inclusion might compromise the composite indicator by introducing bias. Any that are judged to risk compromising the indicator would then be excluded.

Improvements to interpretation

Despite the excellent summary of potential interpretation pitfalls on the BTO website, one issue not discussed: the indicator averages (on a log scale) relative abundances. Thus even if total number of individuals is constant, the indicator can show a significant trend – downwards if common species tend to be increasing while rare species tend to be decreasing, and upwards if the converse is true. The indicator thus reflects trends in both abundance and evenness. Use of additional subsidiary indicators to allow these two components to be separated should be considered.

C6a,b – Insects of the wider countryside (butterflies)

Phil Stephens

Background

This includes C6a (Semi-natural habitat specialists) and C6b (Species of the wider countryside). These indicators draw on data from the same monitoring programmes. Other species groupings are possible within the broader set of monitored species (e.g. the farmland and woodland indices) (see Brereton et al. 2011a) but the underlying methods are the same for all. Consequently, they are reviewed together.

The indices themselves are illustrated below (Fig. 1). C6a is based on 26 species, whilst C6b is based on 24 species. Together, these represent all 50 of the species that are regularly resident in England.

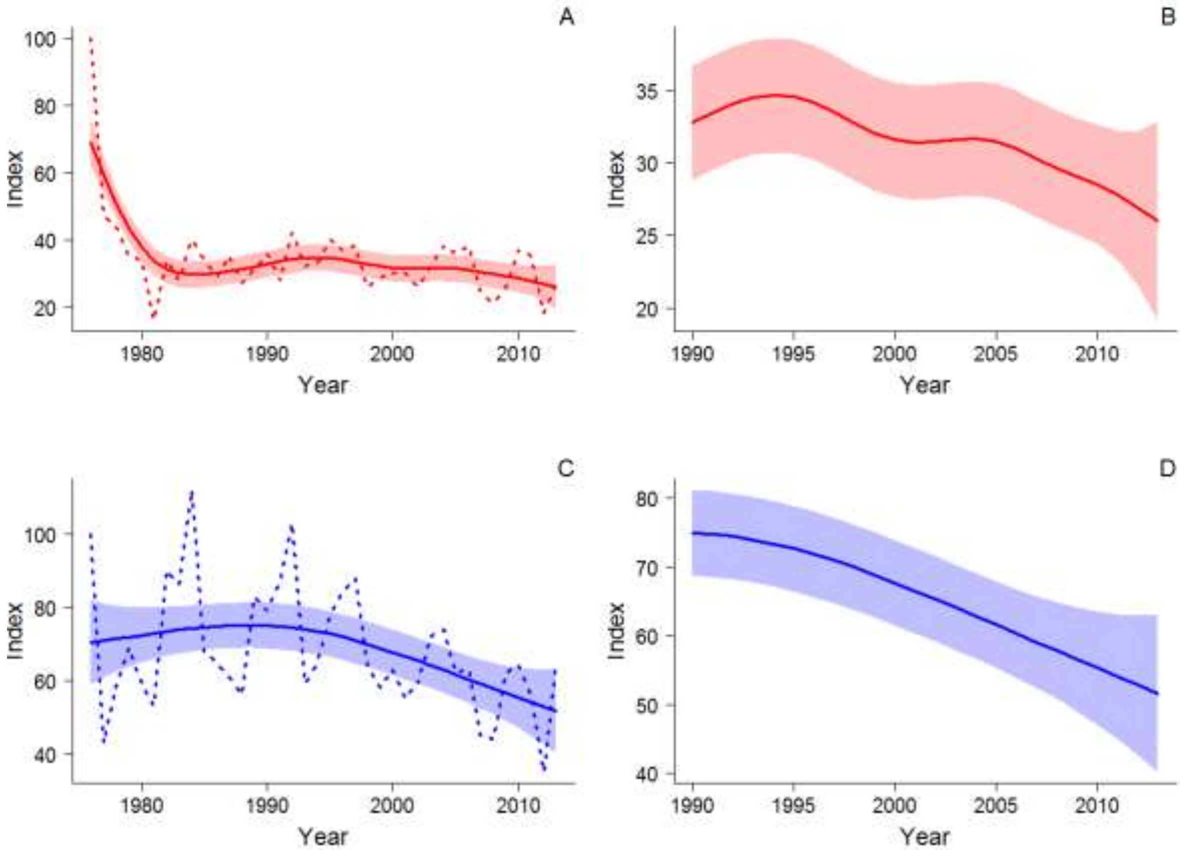


Fig.1. The C6a Semi-natural habitat specialists indicator (A,B) and C6b Species of the wider countryside indicator (C,D). Panels A and C show the unsmoothed trend (broken line) and smoothed trend (solid line) with 95% confidence interval. Abrupt changes in the early years, especially for the habitat specialists, diminish the appearance of recent trends (A,C). However, a focus on the smoothed trends since 1990 shows that both indices are in decline (B,D).

Data Quality

Survey design

Butterfly indicators are based on data from multiple sources, including:

(1) high-intensity (weekly through summer) transects at volunteer-selected sites initially run (since 1976) by the Butterfly Monitoring Scheme (BMS) and merged in 2005 with another set run by Butterfly Conservation to give the UKBMS (e.g. Brereton et al. 2011a);

(2) lower-intensity (less frequent) transects or adult timed counts (to sample fritillary and other colonial habitat specialist species occupying relatively inaccessible habitats) and larval web counts (to monitor several species whose immature stages are generally easier to record); species monitored by these methods include the heath fritillary (transects), high brown fritillary (timed count), marsh fritillary (larval web count) and brown hairstreak (egg count);

(3) the Wider Countryside Butterfly Survey (WCBS), established in 2009, using randomly allocated squares (or those already monitored through the BTO's Breeding Bird Survey), which are surveyed at least twice over the July and August period with visits spaced at least ten days apart (Brereton et al. 2014).

Hereafter, these will be referred to as Methods 1 to 3.

There is extensive documentation on the design of the basic survey (i.e. Method 1) (e.g., Pollard 1977; Pollard & Yates 1993). A lot of work has gone into validating Method 1 (e.g., see Pollard, Hall & Bibby 1986) and many papers reporting the design and its analysis have been peer-reviewed. "Methods to monitor butterfly abundance are well described, extensively tested and scientifically sound" (Brereton et al. 2011a, p140).

Methods 2 (Warren et al. 1984; Lewis and Hurford 1997) and 3 (Brereton et al. 2011b) are also peer-reviewed. All survey methods have been designed by (or had subsequent input from) statistically knowledgeable people (although not uniformly in cooperation with statisticians), published and peer-reviewed.

Where are the surveys done?

The 3 Methods differ in their site choice. Method 1 is based on volunteer-led site choice and is largely ad hoc. A map of over 1700 sites that have contributed to the scheme is given in Brereton et al. (2011a) and reproduced here (Fig. 2A). Clearly, coverage is biased by both human population density and the location of species-rich butterfly habitats. Central southern England is particularly well-represented relative to other areas (Brereton et al. 2002).

It has long been recognised that this leads to an unrepresentative sample. For example, Pollard et al. (1986, p11) noted that collated index values for sites within a region were not representative of the wider areas, because (i) the proportion of nature reserves in the scheme was very large; and (ii) sites were often managed. Some effort has been made to target new transects to remedy under-representation (e.g. Brereton et al. 2002) but, as is clear from Fig. 2, bias remains. It is thought that this bias can be overcome, to some extent, by post-stratification of transects (e.g. van Sway et al. 2013). Within those constraints, the survey is comprehensive (i.e., all observed, emergent adults are counted throughout the flying season).

Method 2 was largely intended to focus on all known sites for heath fritillaries. Warren et al. (1984) report that they surveyed ‘known and potential’ sites for heath fritillaries, having consulted historical records and given known habitat affiliations. Given the conservation status of the species, it is unlikely that sites of occurrence would be unknown. Within those sites, all observed, emergent adults are counted at each visit – but this is not a complete count. By contrast, for other species (such as the marsh fritillary), no comprehensive statement of site selection is readily available. Surveys focus on larval webs in late Summer and may involve complete counts (if the site is small), or transects (in larger sites).

Method 3 is deliberately designed to employ grid squares selected by random stratified sampling (following the BBS). Nevertheless, coverage is dependent on volunteer effort. Hence, coverage remains regionally biased towards southern England (Fig. 2B).

As with other indices based on sites selected by volunteers, or those of known occurrence or high abundance, these methods risk ‘regression to the mean’ and the appearance of decline.

A **B**

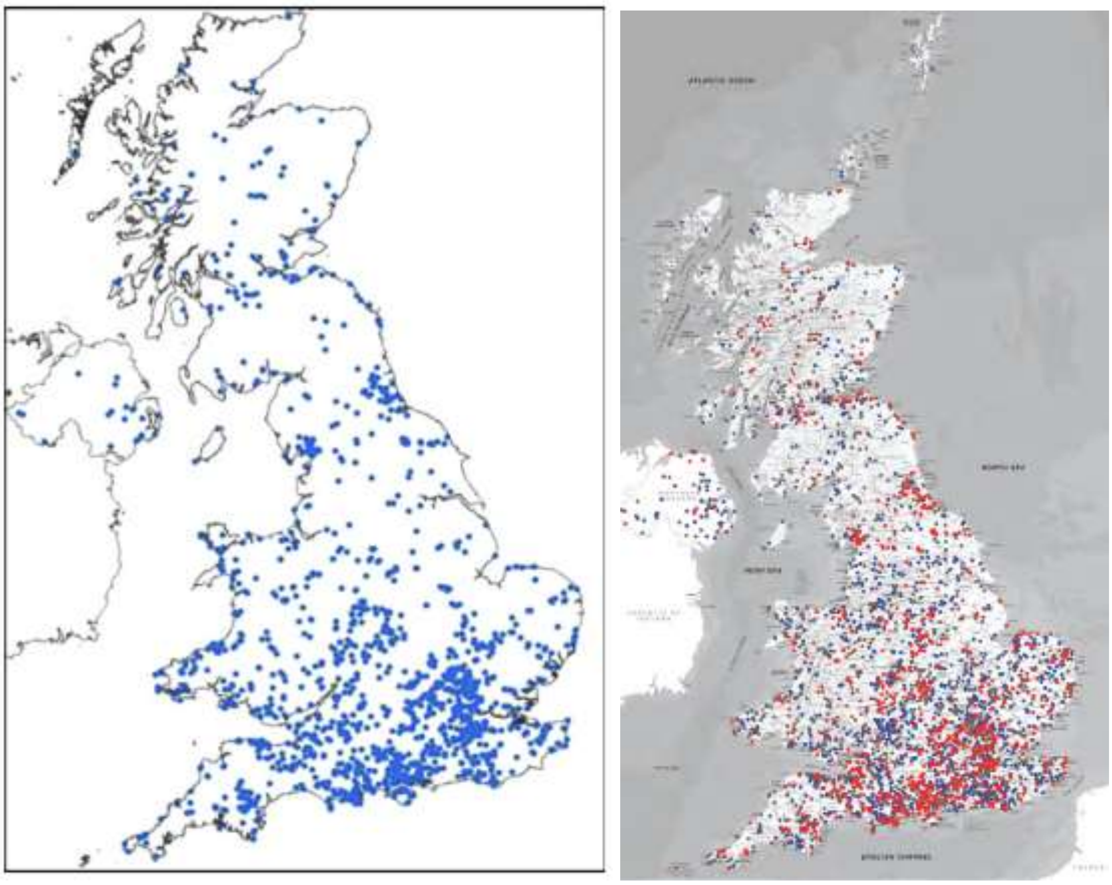


Fig. 2. (A) Locations of 1,700 sites that have contributed to the UK Butterfly Monitoring Scheme (UKBMS) (from Brereton et al. 2011a); (B) UKBMS (red) and WCBS (blue) sites monitored in 2013 (from Brereton et al. 2014).

Changes over time in survey plot location

Sites surveyed using Method 1 grew slowly in number from 35 in 1976 to approximately 120 in 2002 (Brereton et al. 2002). Subsequently, there was rapid growth, such that Brereton et al. (2011a) reported that more than 1000 sites were monitored in 2009. Brereton et al. (2014) estimated a turnover of approximately 7% of sites each year. No equivalent information is available regarding site turnover for Method 2. For Method 3, numbers of sites are still growing but with some turnover; data on the level of turnover undoubtedly exist but are not readily available. Likewise, participants vary across all methods and sites but quantification of participant turnover is not readily accessible. Given the nature of participation, it is likely that – for all methods – turnover is random within the set of available sites. Turnover rates for all Methods, for both sites and participants, are desirable.

Fieldwork methods

Fieldwork methods for all approaches are well documented, both in published literature (see citations above) and in guidelines provided to participants. Training courses for participants are run by Butterfly Conservation but attendance at these is entirely voluntary; information regarding the proportion of volunteers who have completed a training course is not readily accessible. UKBMS recommend that new recorders should be accompanied by the main recorder two or three times before recording a transect on their own (http://www.ukbms.org/Downloads/UKBMS%20G2%20Transect%20field%20guidance%20%20notes.pdf).

Considerable information exists to ensure that surveys are conducted at appropriate times and in appropriate weather. To compensate for missed site visits, and to make the most of data from less intensive methods, a modelling process is now used to impute missing data, account for year differences in seasonal flight periods, and account for site differences in abundance (Dennis et al. 2013). At present, however, this only utilises data collected by the transect methods (Dennis et al. 2013, p639) and ignores site differences in flight periods (*ibid.*, p644). Other, less intensive methods seem vulnerable to abrupt phenological shifts.

For transect methods, detectability has been considered. Variation in detectability has been found to be small compared with variation in true abundance, such that population density estimates from transects are highly correlated with those derived from distance sampling (Isaac et al. 2011). As the technique focuses on relative rather than absolute abundance, this is reassuring.

For transect methods, location of transects is unlikely to be random but the focus is on year to year variation at fixed transect sites. Once established, transects will seldom vary in location. However, if they do change location (e.g., if two fields are merged by removal of a hedgerow and margin along which the transect was conducted), it is unclear how observers should respond. Arguably, the most rigorous response would be to continue to walk the same route but this might be neither appealing to the volunteer, nor appropriate from the landowner's perspective. If transects are routinely in the best locations within the surveyed square, there is a possibility for buffer effects (i.e., increases and decreases in butterfly abundance might be more marked in areas beyond the best habitats).

Data quality assurance

There is no obvious reference to the systematic identification of outliers, so this is probably an *ad hoc* process. However, it might occur and not be reported formally. New recorders' forms should be checked for any anomalies when they return from doing a transect count (http://www.ukbms.org/Downloads/UKBMS%20G2%20Transect%20field%20guidance%20%20notes.pdf).

Rigour of Analytical Method

Trends are noisy and so are smoothed before confidence intervals are calculated (see Fig. 1). Confidence intervals are not given for annual indices (until those have been smoothed). Smoothing, confidence intervals and estimates of temporal change are all calculated using the Trendspotter software (Visser 2004), although there are plans to use GAMs for smoothing, as is now the case for indices of bird abundance. Treatment of the data provided to Trendspotter is unclear. Some limitations arise from the conversion of weekly counts to an annual index. In particular, recent literature makes no reference to annual (e.g. climate- or resource-driven) changes in either activity or longevity, both of which would undermine estimates of inter-annual change (and, if they are also experiencing trends, might undermine longer term trends also). Pollard et al. (1986) recognised these concerns; validation work did not identify activity as problematic but shifts in longevity do not appear to have been studied.

Whether the different survey methodologies can be combined with conviction is unclear. For example, Brereton et al. (2011a, p142) cite Brereton et al. (2002) as showing that transect and non-transect data give similar regional trends and combining those data can increase the precision of trend estimates. However, there is no obvious consideration of this issue in Brereton et al. (2002); it is possible that there are differences in the content of this report between the copy produced by Butterfly Conservation, and that available via the Defra website (the version consulted for this assessment).

Precision and Bias

The butterfly indices are subject to all four forms of bias of interest in the context of these assessments. Geographic bias has already been discussed (see Fig. 2). Post hoc stratification of sites and comparisons between the UKBMS and WCBS can give indications of geographic bias but it remains the case that, when the full data set is used, the index is principally of use for identifying trends across sampled sites. Observation bias, owing to changes in recorder effort over time, is an inherent problem in any programme which relies on volunteer effort (and which seeks to increase participation over time). Raw indices, especially in the early years, might have been strongly subject to this form of bias but minimising the impacts of observation bias is a focus of statistical techniques for smoothing the index. Reporting bias, resulting from selective recording by observers, is obviously a possibility. This could arise especially where rarer species could be mistaken for more common alternatives. Detection bias could also arise via changes in detectability over time. However, note comments under 'Fieldwork methods', above, that suggest that variation in detectability has been found to be small compared with variation in true abundance.

The time series is long, informative and sufficient to provide a good indication of trends at a multi-annual scale of interest to the government and stakeholders.

The methods used to create the multi-species indices are the same as those used to create multispecies trends for common birds and are the state of the art.

Interpretation

For each index, species-specific trends are combined using a geometric mean of relative abundances. Thus, the butterfly indices are composite trends that, as with others reviewed here, are typically interpreted to indicate changes in abundance of the focal species over time. As such (and, again, as with the other composite trends), they are vulnerable to the conflation of effects due to changes in evenness and those due to changes in overall abundance.

The indicators are typically updated each year, which is appropriate and justifiable (although there are some lags in the acquisition of data). Raw indices are noisy, whilst smoothed indices are unresponsive. Short-term changes in trends (of one or only a few years) are unreliable and should not be reported.

The butterfly indicators form part of the evidence base which is synthesised for reporting, to inform a wider policy evaluation of progress towards the Aichi targets. The butterfly indices are considered relevant to, and contribute to reporting on, a range of Aichi Targets. Although they do not map directly to any of those targets, it is clear that the production of the indicators motivates the collection of data that can be subjected to further analyses to cast light on the targets.

Conclusions

We conclude that the quality, clarity and relevance of this indicator could be improved in the following ways:

Improvements to data collection

The switch to stratified random sampling (using the grid identified for the BTO's breeding bird survey) is the most significant aspect of data collection that could be changed. This would have substantial benefits for the utility of the index, greatly enhancing the potential to claim its representativeness. Clearly, with the introduction of the WCBS in 2009, the feasibility of this switch is already under consideration. The specific method of volunteer recruitment is unclear – but clearly a targeted campaign to increase the amount and consistency of monitoring in regions with lower participation will help to reduce the danger of geographic bias (stratification notwithstanding).

Protocols for the choice of transect routes are not completely clear in the literature relating to the butterfly surveys. If, for example, a hedgerow between two fields forms the basis of a transect route and is removed, does monitoring continue to be conducted through the centre of the new, larger field, or is it relocated to another hedgerow? This might be a relatively rare occurrence but is a specific (and extreme) example of how habitat modification would be dealt with. Greater clarity over site-level choices like this is important to interpreting the index.

Improvements to analysis methods

The proposal to move to the use of GAMs for smoothing the index and for identifying year and site effects is welcome. Some useful validation work could be conducted to monitor trends in longevity and their consequences for index values.

Improvements to interpretation

See generic issues relating to the use of composite trends, and the difficulties of differentiating changes in abundance from changes in evenness.

References

Brereton T, Warren M, Stewart K (2002) BD1427: developing a system for assessing the effect of agri-environment schemes on butterfly populations. Butterfly Conservation, Wareham [which, here, is assumed to be the same as the report here:

http://sciencesearch.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=7999]

Brereton, T., Roy, D.B., Middlebrook, I., Botham, M. & Warren, M. (2011a) The development of butterfly indicators in the United Kingdom and assessments in 2010. Journal of Insect Conservation, 15, 139–151.

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

Brereton, T.M., Botham, M.S., Middlebrook, I., Randle, Z. Noble, D.G. & Roy, D.B. 2014. United Kingdom Butterfly Monitoring Scheme report for 2013. Centre for Ecology and Hydrology and Butterfly Conservation.

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

Gregory, R.D., Willis, S.G., Jiguet, F., Voříšek, P., Klvaňová, A., van Strien, A., Huntley, B., Collingham, Y.C., Couvet, D. & Green, R.E. (2009) An indicator of the impact of climatic change on European bird populations. PLoS ONE, 4, e4678.

Isaac, N.J.B., Cruickshanks, K.L., Weddle, A.M., Rowcliffe, J.M., Brereton, T.M., Dennis, R.L.H., Shuker, D.M. & Thomas, C.D. (2011) Distance sampling and the challenge of monitoring butterfly populations. Methods in Ecology and Evolution, 2, 585–594.

Lewis OT, Hurford C (1997) Assessing the status of the marsh fritillary butterfly (*Eurodryas aurinia*): an example from Glamorgan, UK. Journal of Insect Conservation 1, 159–166

Pollard E (1977) A method of assessing changes in the abundance of butterflies. Biological Conservation, 12, 115–134

Pollard, E., Hall, M.L., & Bibby, T.J. (1986) Monitoring the abundance of butterflies 1976-1985 Nature Conservancy Council, Peterborough.

Pollard, E. & Yates, T.J. (1993) Monitoring Butterflies for Ecology and Conservation. Chapman & Hall, London.

Warren MS, Thomas CD, Thomas JA (1984) The status of the heath fritillary butterfly *Mellicta athalia* Rott. in Britain. Biological Conservation 29, 287–305

van Sway & many others (2013) The European Grassland Butterfly Indicator: 1990–2011. EEA Technical Report 11/2013. doi:10.2800/89760

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

C8a - Bats

Ken Norris

Background

This assessment covers C8a – Bats. The indicator represents changes in the population size of eight bat species from 1999 onwards; and is generated by the Bat Conservation Trust based on data collected by the National Bat Monitoring Programme (NBMP). Methodological, analytical and population trend details are reported in Barlow et al. (2015). The composite indicator shows an increasing trend from 1999 to 2012, with a more stable in recent years.

Data Quality

Survey design

NBMP is a citizen science scheme that collects data using four main approaches – roost count, hibernation survey, field survey and waterway survey. Details of survey designs, fieldwork methods and species population trends have recently been published (Barlow et al. 2015).

- Roost counts – these are carried out at summer roosts of seven species located in buildings and other man-made structures.
- Hibernation surveys – these are carried out at hibernation sites, including caves, mines and other underground structures e.g. cellars. All species encountered are counted.
- Field surveys – these collect activity data on four species along transects using bat detectors.
- Waterway survey – this collects activity data for Daubenton's bat along watercourses.

Where are the surveys done?

- Roost counts – at self-selected (i.e. known) sites.
- Hibernation surveys – at self-selected sites.
- Field surveys – based on 1km grid squares selected using a stratified, random approach based on the proportional representation of 40 land classes from the UK Land Cover Map 2000.
- Waterway survey – based on transect centred on an allocated grid reference associated with a watercourse >2m wide and with River Habitat Survey data.

Changes over time in survey locations

Turnover of locations is considerable – the annual number of sites covered can be as low as only 13% of the total sites covered throughout the time series.

Fieldwork methods

All surveys are restricted to suitable weather conditions. Specific methods for the different components are as follows:

- Roost counts – emergence counts, starting at 15mins prior to sunset or at sunset depending on the species, are undertaken by volunteers at two dates between May and July.
- Hibernation surveys – two daytime visits are made by groups of surveyors (one in January, one in February). Surveyors follow a standard route through the site searching open

locations and crevices. Some counts done outside (December or March) standard period (i.e. some site-dependent variation in protocols).

- Field survey – 3km triangular-shaped transect mapped out within each 1km square and split into 12 approximately equal sections. Surveyors walk each section of the transect with a heterodyne bat detector tuned to 25 kHz and count the number of bat passes (a sequence of two or more echolocation calls heard as the bat flies past the detector and separated from a previous pass by at least 1 second). At the end of each section, surveyors stop and conduct a 2-min point count with the detector tuned to 50 kHz and the number of bat passes counted. Transect walks are conducted 20mins after sunset on two dates in July separated by at least 5 days.
- Waterway survey – a 1km transect centred on the grid reference is divided into ten locations evenly spaced along it. Surveyors complete a 4 min point count at each location and count the number of bat passes using a bat detector tuned to 35 kHz, or they record activity as continuous at locations in which activity is very high (i.e. where it is difficult to distinguish individual passes).

Data quality assurance

Bat identification is a major issue since this is a volunteer survey. NBMP provides training resources to ensure volunteers have the relevant skills to undertake the surveys, but no data are presented on identification accuracy. Training includes bat detector workshops, online training tutorials, field notes on identification and video demonstrations. No details on any post-survey data checking protocols are available.

Rigour of Analytical Method

Trends are analysed using a log-linear GLM with Poisson errors, which are fitted to the count data from each survey type and species. Models include a site term and the time trend is then modelled using a GAM framework to produce a smoothed trend. The basic analytical framework is the same one used for breeding birds. The geographical distribution of volunteers is uneven so counts are weighted in certain circumstances to account for this. Maximum counts per year are used in the models for roost data for improved precision (implies individual counts were variable and hence not very repeatable). Binomial models of the proportion of point counts or transect sections in each survey where the species was observed are used to analyse the trends based on activity data (field and waterway surveys). This is due to problems of over-dispersion in the activity data arising because individual bats may repeatedly fly past the detector. More complex GAMs are fitted to the field survey data that include covariates for microphone type and sensitivity range to account for the fact that the type of bat detectors used varies across the survey in both space and time. Overall, the analytical framework is reasonably rigorous and draws heavily on the methods used to analyse bird population trends. Obvious sources of bias in the data have been identified and analyses modified to address these.

Precision and Bias

Bootstrapped confidence intervals are calculated for each smoothed trend (i.e. survey type and species) by creating new datasets based on resampling with replacement sites from the original dataset. At least 400 bootstrap samples were created for each trend model. Power analyses were undertaken to estimate the number of survey years required to detect specific rates of decline based on the field survey data. For roost counts and hibernation survey data, power to detect a decline was assessed relative to the width of the confidence interval. Data appeared adequate to detect severe (i.e. 50% over 25 years) but not moderate (i.e. 25% over 25 years) declines. Sources of bias are likely to be significant. Observation bias arises because of the inclusion of counts from known roost and hibernation sites. This bias almost certainly explains why trends within species differ in sign and magnitude across survey methods – only one of eight species shows consistency in both the direction and rate of change across surveys. Reporting bias is possible but difficult to assess at present. Observers are given training in identification but repeatability between observers has not been investigated. Detection bias arises because of changes in the types of bat detector used over time. Geographical bias arises because of an uneven distribution of observers across GB. Detection and geographical bias are addressed in the data analysis.

Interpretation

The indicator is a composite index derived from all sources of survey data across the eight species. Each species is given equal weighting and the annual index for a particular year is the geometric mean for that year. The indicator is a proxy of (breeding) population size and no data exists linking the survey data with direct counts of breeding bats.

Conclusions

We conclude that the quality, clarity and relevance of this indicator could be improved in the following ways:

Improvements to data collection

It would be valuable to assess reporting bias in the various surveys by having a sample of sites counted by multiple, independent observers.

Improvements to analysis methods

None.

Improvements to interpretation

Consider reporting an indicator based only of the field and waterway surveys to reduce the observation bias inherent in the counts from known roost and hibernation sites. Use the field survey data to assess both changes in abundance and evenness in abundance trends across species to improve interpretation of the indicator.

References

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

D1- Status of Pollinating Insects

Ken Norris

Background

This assessment covers D1 – Status of Pollinating Insects. The indicator represents changes in the occupancy of sites by 216 bee species since 1980; species selection is based on data availability. It is generated using unstructured biological records by the Centre for Ecology & Hydrology. No assessment of trends in the indicator have yet been made pending the development and application of appropriate analytical methods.

Data Quality

Survey design

There is no standard sampling design. The data are based on unstructured volunteer records.

Where are the surveys done?

Nationally but no map of observer coverage is currently available.

Changes over time in survey locations

Probably considerable but no data are available on this.

Fieldwork methods

There are no standard methods.

Data quality assurance

Records are usually checked for accuracy by a local records co-ordinator, but otherwise none.

Rigour of Analytical Method

Trends are analysed using novel Bayesian occupancy models for estimating species occurrence in the presence of imperfect detection (Isaac *et al.* 2014). The approach uses two coupled sub-models – an occupancy sub-model and a detection sub-model. The model is used to estimate the annual proportion of sites occupied for each species, then a linear model is fitted to these annual occupancy estimates to assess the trend. These trends are converted into an index and the indicator calculated as the geometric mean of the index across all species (equal weighting).

Precision and Bias

Confidence intervals for the indicator are calculated using bootstrapping based on 10,000 datasets generated by resampling bee species with replacement and recalculating the geometric mean for each dataset. Potential biases are significant and include observation (changes in effort over time), reporting (incomplete and selective reporting by observers), detection and geographical bias.

Interpretation

Recent work suggests that the occupancy models used produce robust trends in the face of observation, reporting and detection bias (Isaac *et al.* 2015; van Strien *et al.* 2013). Geographical bias in the indicator remains to be investigated. The indicator is actually a proxy measure for pollination services and there are no data linking the indicator to service delivery.

Conclusions

We conclude that the quality, clarity and relevance of this indicator could be improved in the following ways:

Improvements to data collection

Acquire estimates of sampling intensity at the point of data collection (Isaac *et al.* 2015).

Improvements to analysis methods

Occupancy models seem to provide a promising approach for dealing with the various biases in unstructured survey data, but would benefit from further testing using simulated and independently derived survey data (e.g. van Strein *et al.* 2013).

Improvements to interpretation

The indicator relates to pollination services. Pollination varies in relation to the functional importance of particular pollinator species and their abundance. At present the indicator does not consider these issues. The functional importance of different species can be addressed either by selection of species to include in the indicator or weightings applied to species when estimating occupancy trends. The implications of this decision need to be explored. Further work is required on the relationships between occupancy, abundance and pollination to better define functionally important levels of occupancy below which pollination services degrade.

References

A short technical paper on the indicator analysis is available online - http://jncc.defra.gov.uk/page-6851.

Isaac, N.J.B., van Strien, A.J., August, T.A., de Zeeuw, M.P. & Roy, D.B. (2014) Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution*, **5**, 1052–1060.

Van Strien, A.J., van Swaay, C.A.M. & 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.

General framework of good practice

The structured reviews of the indicators reveal a range of cross-cutting issues that affect the quality of indicators and hence the reliability of status and trend estimates based upon them. Here we synthesize these issues and identify ways to help improve the quality and reliability of the indicators. This synthesis revolves around four key issues:

1. Species selection
2. Data quality
3. Trend analysis, including bias reduction and estimates of uncertainty
4. Interpretation

We deal with each of these in turn.

Species selection

Ideally, an indicator should be based on a sub-set of species that adequately represents the wider community of which it is part. In this case, status and trends in the indicator are likely to reflect status and trends in the wider community. In practice, the inclusion of species in an indicator is often subjective or simply reflects the availability of data. As a result, status and trends are likely to be biased to some extent, but the sign and magnitude of the bias remains uncertain. To address this issue, we identify the following improvements:

- For a few indicators (e.g. C5a – farmland birds), data are available on a wider range of species than the sub-set of species included in the indicator. In such cases, it would make sense for species selection to be done objectively rather than subjectively. Objective methods are available (e.g. Butler et al. 2012; Wade et al. 2014), and others could be developed. Furthermore, even if a current indicator uses all available data, objective methods could be used to define an indicator sub-set that better represents the wider community. Inevitably, there may be trade-offs between uncertainty in status and trends and the number of species included in the indicator, but such trade-offs can and should be explored.
- For certain indicators representativeness within and /or between taxonomic groups is so poor that we question the utility of the indicator (e.g. C4a – status of priority species – relative abundance).
- At the very least, a statement about representativeness should be drafted for each indicator, and should include an assessment of the number of species included in the indicator compared with the number in the wider community, and acknowledgement of any known biases in the indicator sub-set (e.g. species in the indicator are known to be declining more rapidly than species in the wider community). For example, this is stated for C4b – status of priority species – frequency of occurrence - insects.

Data Quality

Broadly, the data underlying the indicators are of two types – count data providing trends in relative abundance, and occupancy data from unstructured biological records. We discuss these two data types separately:

Count data

In many cases, a complete survey of all known sites in which a particular community occurs is impossible for logistical reasons. As a result, a sample survey is frequently used to generate the count data underlying a particular indicator. Ideally, a stratified, random sampling design should be used to do this because it minimizes bias in the data. Although a number of indicators are based upon such a design, a number of others rely, at least in part, on sites selected by volunteers. For example, known roost and hibernation sites are surveyed for bats (C8a) and volunteers select sites to survey for butterflies (C6). Since it is typical for volunteers to concentrate activity at sites in which animals are likely to be relatively abundant, this approach introduces significant observation bias into the data because such sites are likely to show much greater negative trends than the population as a whole (a regression to the mean problem). This bias is apparent, for example, in the bat indicator in which there is little consistency within species in the sign and magnitude of trends estimated using volunteer-selected and random sites (see Barlow et al. 2013). As a result, we suggest that:

- Attempts should be made to phase out volunteer-selected sites and replace them with a stratified, random sampling design. This transition has successfully been made from the Common Birds Census (CBC) to the Breeding Bird Survey (BBS), but a similar approach needs to be adopted across indicators.
- Indicators that currently include both data from volunteer-selected and random sites should report separate status and trend estimates from these different sampling designs, and place greater weight on the interpretation and use of trends based on random sites.

Even if a stratified, random sampling design is used significant bias remains possible. Reporting bias can occur if particular species are consistently misidentified by observers, or if particular abundances are consistently under-estimated or over-estimated by observers. Although volunteer-based surveys frequently provide training to observers, we found very little evidence relating to the effect of such training on reporting bias. As a result, we suggest that:

- Where reporting bias is considered likely due to difficulties in identifying particular species or estimating particular abundances, studies are conducted to assess reporting bias and the extent to which it is reduced by training programmes.

Detection bias can occur if detection rates vary over time. In this case, trends in the indicator may reflect changes in detection rather than genuine changes in relative abundance. We found no strong evidence across indicators of this issue, although we note that it is difficult to quantify without adequate data. There is some evidence from the Breeding Bird Survey (BBS) that detection rates can change over time and that trend estimates can differ substantially as a result of the detection bias (Newsom et al. 2013). As a result, we suggest that:

- Where possible, analyses should be conducted to check for changes in detection and trend analyses modified accordingly.

We note that geographical bias is an issue for all volunteer-based surveys due to spatial variation in human population densities across the UK. Without resources to target additional survey effort this bias needs to be addressed in the trend analysis (see below).

Occupancy data

Indicators are increasingly using data from unstructured biological records (e.g. B6 – Pressure from invasive species, D1c – Status of pollinating insects). It is well known that such data routinely contain significant observation, reporting, detection and geographical bias (Van Strien et al. 2013; Isaac et al. 2014). As a result, we suggest that:

- Trends based on raw data from unstructured biological records are not used directly and that only trends estimated from statistical models (e.g. occupancy models) that reduce the biases are used, such as those being developed and used for certain indicators (e.g. D1c – Status of pollinating insects).

Trend analysis

For trend analysis, abundance is usually expressed as an index relative to a baseline year towards the beginning of the time series. For certain time series, the first year of data is anomalous when compared with index values for subsequent years. Using such a year as a baseline unnecessarily complicates interpretation of trends in the index. As a result, we suggest:

- Either a suitable baseline year is selected on the basis of data availability (e.g. Barlow et al. 2013) or by smoothing the index values using a method that is not sensitive to index values at either end of the time series (see also following paragraph).

Our review revealed a range of approaches to trend analysis, including interpreting the raw data directly without any formal analysis. As noted above, the data underlying the indicators are of two broad types – count data providing trends in relative abundance, and occupancy data from unstructured biological records. Statistical modelling frameworks are available for both types of data. Trends in count data can be analysed using generalised additive models (Fewster et al. 2000; Barlow et al. 2013), within which smoothed trends can be fitted to the data. This framework allows the inclusion of factors that reflect sources of bias or noise in the data (e.g. Barlow et al. 2013). Trends in occupancy can be analysed using occupancy models (Strien et al. 2013), which seem to be a promising approach for addressing the biases inherent in unstructured data. We suggest the adoption of a common trend analysis framework for indicators based on the different data types that include:

- The use of GAMs for trend analysis based on count data including a structured approach to the elimination of likely biases in the data (following the approaches adopted by Fewster et al. 2000 and Barlow et al. 2013).
- The use of occupancy models for trend analysis based on occupancy data from unstructured biological records (following the approach outlined in Strien et al. 2013). We advise against the use of raw data in this case because of the inherent biases. We also note that although occupancy models appear promising, further work is required to assess model assumptions and performance across a wider range of conditions.

Confidence intervals in the estimated trends are usually produced to provide a measure of uncertainty. This is typically done using bootstrapping based on the resampling of either sites or species. Our review revealed no clear justification for the approach to bootstrapping associated with the different indicators. Bootstrapping based on sites provides a reasonable measure of uncertainty if the data come from a random sampling survey or if surveyed sites are representative of all of the sites that could have been surveyed. Even if there is geographic bias in the sample, post-stratification can be used to improve representation. Bootstrapping based on species assumes that the sub-set of species of species in the indicator is representative of the species in the wider community, which is frequently not the case (see above). Furthermore, variation in trend between species is likely to be greater than between sites, so uncertainty in the trend is likely to increase if bootstrapping is based on species rather than sites. As a result, we suggest that:

- Bootstrapped confidence intervals in estimated trends should be generated by re-sampling sites wherever possible.

Interpretation

Most indicators are produced by calculating the geometric mean of the relative abundance index across species and surveys. Trends in the geometric mean are then frequently interpreted as if the trends represent the absolute change in abundance across species. This may be the case if all species in the indicator are changing at the same rate, but this is unlikely to be true. As a result, trends in the geometric mean can arise due to changes in the evenness of trends across species (Buckland et al. 2011). Where trends are based on a robust survey design such as BBS, it is possible to test for changes in evenness across species and hence aid interpretation of trends in the geometric mean (Buckland et al. 2011). As a result, we suggest that:

- Evenness statistics are routinely calculated and reported alongside trends in the geometric mean index to aid interpretation wherever possible.

References

Buckland, S.T., Studeny, A.C., Macgurran, A.E., Illian, J.B. & Newson, S.E. (2011) The geometric mean of relative abundance indices: a biodiversity measure with a difference. Ecosphere 2, 1-15.

Butler, S.J., Freckleton, R.P., Renwick, A.R. & Norris, K. (2012) An objective, niche-based approach to indicator species selection. Methods in Ecology & Evolution 3, 317-326.

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

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

Isaac, N.J.B., van Strien, A.J., August, T.A., de Zeeuw, M.P. & Roy, D.B. (2014) Statistics for citizen science: extracting signals of change from noisy ecological data. *Methods in Ecology and Evolution*, **5**, 1052–1060.

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

Van Strien, A.J., van Swaay, C.A.M. & 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.

Wade, A.S.I., Barov, B., Burfield, I.J., Gregory, R.D., Norris, K., Vorisek, P., Wu, T. & Butler, S.J. (2014) A niche-based framework to assess current monitoring of European forest birds and guide indicator species' selection. PLoS One 9, e97217.