British Trust for Ornithology

BTO Research Report No. 257

Bird Indicators of Sustainability for the Water Industry

Authors

G.M. Siriwardena, J.J.D. Greenwood & N.A. Clark

June 2001

A report by the British Trust for Ornithology under contract to Northumbrian Water plc.

British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU, UK.
Registered Charity No. 216652

CONTENTS

I. INTRODUCTION

In 1998, the UK government (specifically the Department of the Environment, Transport and the Regions (DETR)) funded work by the British Trust for Ornithology (BTO) and Royal Society for the Protection of Birds (RSPB) to produce one of several indicators of sustainable development in the UK from existing and ongoing bird surveys (Gregory *et al.* 1999). Birds were chosen to signify biodiversity and environmental health in general because they are easily the best-monitored taxon in the UK. The resulting indicator became one of 14 measures originally publicized as tools for the future monitoring of the sustainability of development.

The idea that key aspects of national biodiversity can be summarized effectively in a single index number, such that future changes will be easily detected and easily communicated to a wide audience, is an attractive one. It also leads naturally to the suggestion that regional and habitat-specific information can be collated in a similar way to assist with management decisions at smaller spatial scales. Indices such as this will never provide a complete assessment of biodiversity or of the health of the environment, but can provide useful pointers given that they are constructed with care and interpreted appropriately.

This report describes exploratory analyses of existing UK bird survey data investigating how indicators of sustainability might be developed from bird census information for one sub-set of habitats for which particularly good summer and winter data are available: wetlands. The work was commissioned by Northumbrian Water and focused both on water bodies owned by the water industry and on wetlands in general, nationally. We have produced indicators for three broad types of wetland (also sub-divided as shown in parentheses):

1) still waters (natural and man-made, reservoirs and gravel pits);
2) linear waterways (large and small);
3) reedbeds, water meadows and other damp sites (separately).

In addition to these, national, indicators, we have investigated regional sub-divisions of the available information, concentrating on the constituent countries of the UK and the English Regional Development Agency (RDA) regions (for several water companies, including Northumbrian Water, the appropriate RDA region is approximately equivalent to the company's catchment area). RDA regions are becoming a key administrative division in the government of the UK and have responsibilities in ensuring the sustainability of development, and they form the geographical basis for new, all-habitats sub-divisions of the headline wild bird indicator (Robinson *et al.* 2000).

We have followed the rationale and methods for the headline indicators project (Gregory *et al.* 1999) because this approach has been accepted by the government and adopted by other indicator projects (e.g. Robinson *et al.* 2000). A number of important issues arise from a critical review of the headline indicators work, which shows that great improvements to the methods used could easily be made. We have attempted to address those issues that are methodological by adapting the approach taken, whilst retaining the overall approach used previously; we consider further, more general issues in the Discussion.

II. METHODS

1. SEASONS

A fundamental consideration in developing indicators of bird abundance in wetland habitats is that very different populations are present at different times of year. As in other habitats, many breeding species are summer migrants (examples include Garganey *Anas querquedula* and Reed Warbler *Acrocephalus scirpaceus*), but there is also a large influx of more northerly breeding birds in winter. The species and population differences between the birds using wetlands in summer and winter mean that attempting to combine monitoring data from the two seasons would be unwise. As a result, we have treated summer and winter separately and produced indicators accordingly.

The large populations wintering in the country means that the UK holds internationally important numbers of many wetland species, particularly waterfowl and waders outside the breeding season. In addition to the potential value of these wetland bird populations as a tool for monitoring their environment (which is equally high for breeding birds), it is therefore also important, conversely, to monitor the winter characteristics of wetland habitats with respect to their importance for the birds.

In addition, UK wetlands are important as migratory stopover sites for many species and they hold large numbers of waterfowl while they are moulting and flightless. Although these periods could be critical for the species concerned, they are monitored less well than the breeding season and mid-winter. The use of wetland sites by birds during moult or migration was therefore not considered in the generation of indicator indices in this study.

2. SOURCES OF DATA

Several long-running and ongoing survey schemes that draw on the efforts of volunteer fieldworkers provide information on birds in wetlands, each concentrating on different habitat types and sets of species. We have used each available data set as described below, aiming where possible for a 25-year run of data (1974-1998):

2.1 BTO/WWT/RSPB/JNCC Wetland Bird Survey (WeBS)

In combination with its predecessors, the Wildfowl and Wetlands Trust (WWT) National Wildfowl Counts and the British Trust for Ornithology (BTO) Birds of Estuaries Enquiry, WeBS has provided monitoring information for wetland birds in Britain for over 50 years. WeBS aims primarily to monitor all non-breeding waterfowl populations in the Britain and covers both a sample of inland sites and almost all of Britain's estuaries. WeBS provides near-complete counts of the populations of many coastal, wintering species, but only partial coverage of the birds of inland wetlands. Routine indexing from WeBS data deals only with wintering birds (counts from September to March), but a subset of survey sites is also counted in summer. Full details of the methods used and the background of the survey can be found in WeBS annual reports (e.g. Cranswick *et al.* 1999).

We have used WeBS data to produce population indices for as many species as possible for British inland still waters (natural and man-made) in both summer and winter. We have also produced indices specific to reservoirs only, to gravel/sand pits only and to pits and reservoirs only, using the standard classification of inland WeBS sites as lake, reservoir or pit. Winter indices could be calculated from data from 1974-1998 for most species, but four were not counted in early years of the survey (see Table 1 for details). In summer, sufficient sites to allow indexing have been counted only since 1994. The methods used here could be extended to cover Ireland via the Irish Wetland Bird Survey (IWeBS), but we have not done so as yet.

2.2 BTO Common Birds Census (CBC)

The CBC has been run by the BTO since the early 1960s (part-financed by the JNCC), monitoring breeding bird populations through a territory mapping method conducted over annual sets of 12 visits to defined census plots. Plot turnover has been considerable since the inception of the survey, but without significant change in geographical or habitat coverage. Geographically, the survey is effectively restricted to lowland Britain because plots are concentrated there (coverage of Ireland is negligible). Unlike WeBS, but in common with the other schemes described below, the CBC indexes national populations through sampling rather than attempted complete counts. CBC methods are described in detail in Marchant *et al.* (1990).

The CBC primarily indexes bird populations in woodland and farmland habitats. However, the habitat in some current and past census plots (designated “special”) has not fallen into either of these categories and many of these plots, as well as some mostly consisting of woodland or farmland, feature wetland habitats. We selected CBC plots incorporating wetland areas by examining the surveyors’ original habitat descriptions. This process allowed population indices to be calculated covering all still waters, reservoirs/gravel pits and linear waterways for 1974-1998.

2.3 BTO/RSPB/JNCC Breeding Bird Survey (BBS)

The BBS (which is run by the BTO) was introduced in 1994 as an eventual replacement for the CBC that would offer equivalent precision with fewer geographical biases and more complete habitat coverage, while allowing more species to be indexed and employing more rigorous statistical methods. The BBS employs sampling units of 1km squares of the national grid, which are surveyed via counts along line transects (on two visits in spring/early summer) and via detailed habitat recording. These squares are selected as a random sample of the UK, stratified by observer density. (The Republic of Ireland is not covered by the BBS, but a sister survey with almost identical methods, the Common Bird Survey (CBS), has recently been initiated and our approach could readily be extended to it.) Birds are recorded in three distance bands (within 25m of the transect line, between 25 and 100m from the transect and between 100 and 200m from the transect) or as in flight. A total transect length of 2km is walked in each survey square and both birds and habitat are recorded separately in each of ten 200m transect sections. The rationale for and development of the BBS, along with a number of evaluation studies, are described in several papers and BTO Research Reports (Field & Gregory 1998, 1999; Gregory & Baillie in press, Gregory *et al*. in press).

The habitat-coding scheme used for the BBS (Crick 1992) is comprehensive, allowing count data for a wide range of wetland habitats to be extracted by 200m transect section. We have used these codes to select transect sections running near still waters, linear waterways and other damp habitats. The details of the codes used to indicate the presence of each habitat are given in Table II.4.1.

Table II.4.1. Habitat codes used to indicate the presence of wet habitats in BBS survey squares. Code letters represent level 1 habitat codes and numbers level 2 codes (Crick 1992): various subdivisions of level 1 habitats C (semi-natural grassland and marsh), D (heathland and bogs) and G (Water Bodies (freshwater)) were used. Note that, because the habitat around each 200m transect section can be described by two combinations of level 1 and level 2 codes, our habitat classifications were not exclusive and some counts may have contributed to the indices being based on more than one habitat type. This is most likely to have occurred if a transect section fell along the division between two of the habitat types we consider or if two habitats were present (such as a river running into a lake). Such situations are unlikely to be common in the data, especially since selections of data were made by transect section rather than by 1×1km square.

Habitat Category	***Codes used***
All still waters	G 1-5 (pond (<50m^2), small water-body (50-450m^2), lake/unlined reservoir, lined reservoir and gravel pit)
Linear waterways	G 6-10 (stream, river, ditch, small canal and large canal)
Large linear waterways	G 7 & 10 (river and large canal).
Small linear waterways	G 6, 8 & 9 (stream, ditch and small canal)
Reedbeds	C 7 (reed swamp)
Water meadows	C 6 (water-meadow/grazing marsh)
Wet heaths	D 2 (“wet heath”)
Bogs	D 4 (“bog”)

2.4 BTO Constant Effort Sites (CES) ringing scheme

This scheme (part-financed by the JNCC) comprises several hundred sites where birds are trapped in mist nets for ringing with a constant trapping effort (i.e. net length and location and time spent trapping). The CES scheme was conceived primarily to provide information on changing annual survival rates through capture-mark-recapture analyses, but the constant effort trapping also allows the estimation of abundance (Peach *et al*. 1996, 1998). The scheme began in 1983 and, although some site turnover has occurred, many sites have been operated consistently for long runs of years. Most sites are located in scrub and reedbed habitats. Although there are a few sites in Northern Ireland, the vast majority are in Britain. CES methods are described in more detail in Baillie *et al.* (1986) and Peach *et al*. (1996, 1998).

The primary value of the CES scheme in the present context is that it monitors different habitats to most of the other schemes considered, particularly reedbeds and wet scrub. We have used reedbed and wet scrub data from the CES scheme both together and separately to produce indicator information. Enough wet scrub CE sites have been operated to allow indices for this habitat (as well as the combined index) to be run from 1983 onwards, but indices for reedbeds could only be calculated from 1986 onwards.

2.5 BTO Waterways Birds Survey (WBS)

The WBS was developed to provide monitoring information on breeding birds along linear waterways, habitats that are not covered well by the CBC. The survey has been operated since 1974 using territory-mapping methods similar to those used for the CBC. Plot coverage incorporates both natural and man-made linear waterways; as with the CBC, plots are selected by observers and therefore do not represent a random sample of rivers and canals across the UK. The geographical biases in the WBS data set are less severe than those in the CBC one, but coverage of Northern Ireland is poor. WBS methods are described in detail in Marchant *et al.* (1990). We have used the complete WBS database for the species monitored regularly by the scheme to construct indicators for linear waterways. The WBS habitat information did not allow the data to be sub-divided into large and small linear waterways; most sites are based around rivers or canals.

2.6 Data sources not exploited in this project

The headline indicators project also used several further sources of breeding season information, notably: percentage 10km square occupancy from the two Breeding Bird Atlases (Sharrock 1976, Gibbons *et al*. 1993), estimates of the breeding numbers of rare species published by the Rare Breeding Birds Panel (RBBP: e.g. Ogilvie *et al.* 1999) and various single species surveys. Each of these sources of information is either weaker or less well-suited to the environmental monitoring that is implicit in an “indicator” than those discussed above: to include them would be to devalue the data with which they are combined. We avoided using the RBBP data for five reasons: (i) some species’ RBBP estimates can be very imprecise; (ii) stochasticity could cause large percentage changes in numbers; (iii) rarity means that populations are unlikely to reflect habitat quality at large geographical scales (local environments at the higher end of a range of habitat qualities might be monitored adequately, but not the wider countryside); (iv) RBBP estimates are not habitat-specific; (v) species-specific management for species such as Bittern *Botaurus stellaris* would tend to blur relationships between features of the environment and numbers at the national scale (changes in numbers are more likely to reflect the management of reserves than changes in the wider countryside). Note that species with fewer than 500 pairs breeding in the UK were excluded from the indices that were adopted as final, headline indicators by the government (Gregory *et al*. 1999).

The treatment of Breeding Bird Atlas data in the headline indicators project involved drawing a straight line between the estimates of the number of 10km squares occupied by a species from each of the two atlases (in 1968-1972 and 1988-1991) and extrapolating up to the present day. Both the use of the atlas presence/absence data as a measure of abundance (with which they are unlikely to be closely correlated) and the treatment of the data in this way are biologically unsatisfactory and statistically invalid so we have avoided them here. BTO policy is now in agreement with this position and a revision of the protocol for the production of the headline indicator that omits Breeding Atlas data is being considered by the joint DETR, RSPB and BTO committee that oversees the production of the index.

Most single species surveys suffer from the same lack of annual information as afflicts atlas data, so would require some kind of interpolation and extrapolation to generate annual indices over long time-series. We have therefore also avoided using these sources of data.

3. SPECIES

We have followed the headline indicators project by using an inclusive approach to the selection of species for each of the indices we have produced. Thus, where survey data can be linked reliably to wetland habitats (all schemes except the CBC), we have considered all native species for inclusion in the final indicator and excluded only those for which statistically valid population indices could not be calculated. This approach meant that *a priori* rules as to threshold levels of rarity above which species would not be considered (as used in the headline indicators project to exclude species too rare to be representative of a significant part of the wider environment: see Discussion) were not required. The methods used for determining the reliability of species' indices are described under "Analyses" below.

An adapted species selection rule was used for CBC data. Although we only selected survey plots that included particular wetland features, most plots selected will also have incorporated substantial areas of habitat not associated with wetland. From the BTO's computerized data archives, it was impossible to determine how many of the birds recorded on a given plot came from a given habitat type. We therefore restricted the species to be considered in analyses of CBC data to those commonly associated with wetlands, thus excluding (as far as possible) the influence of the non-wetland components of each plot. Note that some typically "wetland" species are also found commonly in non-wetland habitats (e.g. Sedge Warbler *Acrocephalus schoenobaenus*, Reed Bunting *Emberiza schoeniclus*). The abundance of these species will have been affected by environmental influences not related to conditions in wetlands as well as those in wetlands themselves, so our attempt to concentrate on wetlands via species selection will probably have worked less well for them than for species like Mallard *Anas platyrhynchos*.

During the planning of this project, a different approach to species selection to that described above was considered, based on the species used as wetland species in the headline indicators project, with additional lists of other wetland species and of all other species found on the survey sites used. We did not follow this scheme explicitly because we had already elected not to use all the same sources of data as were used in the headline project and because we wanted to avoid arbitrary *a priori* decisions as far as possible. Nevertheless, our survey-scheme specific indices will effectively have divided species along similar lines: habitat-specific surveys such as WeBS, WBS and CES clearly only cover birds in wetland habitats (and often birds which are specialists in such habitats) while general surveys such as BBS and CBC cover a wider species range.

Following the headline indicator approach, we calculated species-specific indices for introduced species and excluded them from the final indicators. (Note that some species have both native and feral populations, but only one of these is monitored by the schemes considered here: an example is Barnacle Goose *Branta leucopsis*.) It would be straightforward to add introduced species to our multi-species indices if it were considered desirable. Decisions such as this about the species to include (and such as whether or not to limit analyses to habitat specialists) raise several issues fundamental to the design of indicators: these are considered in the Discussion.

Monitoring data for most wetland species were therefore available (with varying geographical biases and precision) from several different schemes. The coverage of each species considered with respect to scheme and season is summarized in Table II.3.1.

Table II.3.1. Survey scheme-specific coverage of the populations of all species included in final indicator indices. "✓" = covered by scheme with sufficiently large sample sizes for reliable indexing across the broadest habitat divisions considered (each habitat sub-division then included a sub-set of the species shown here: Table II.4.1). "Still" refers to all still waters, "linear" to all linear waterways and "damp" to all damp sites (reedbeds, wet heaths, bogs and water meadows). Introduced species are shown in Italics; they were omitted from the index calculations. Species that are closely associated with wetlands were included where they were found sufficiently commonly by schemes monitoring wetland habitats (see text for details).

	Survey Scheme								
Species	***Winter WeBS***	***Summer WeBS***	***CBC***		***WBS***	***BBS***			***CES***
			Still	Linear		Still	Linear	Damp	
Barnacle Goose		✓							
Bewick's Swan	✓								
Blackbird						✓	✓	✓	✓
Blackcap						✓	✓	✓	✓
Black-headed Gull						✓	✓		
Blue Tit						✓	✓	✓	✓
Bullfinch						✓	✓		✓
Buzzard							✓		
Canada Goose	✓	✓	✓		✓	✓	✓		
Chaffinch						✓	✓	✓	✓
Chiffchaff						✓	✓	✓	✓
Coal Tit						✓	✓	✓	
Collared Dove						✓	✓		
Common Sandpiper					✓	✓	✓		
Coot[1]	✓	✓	✓	✓	✓	✓	✓		
Cormorant[1]	✓	✓							
Crow						✓	✓	✓	
Cuckoo						✓	✓	✓	
Curlew					✓		✓	✓	
Dipper				✓	✓		✓		
Dunnock							✓	✓	✓
European Whitefront	✓								
Farmyard Goose		✓							
Feral Pigeon							✓		
Gadwall	✓	✓							
Garden Warbler						✓	✓		✓
Goldcrest						✓	✓	✓	
Goldeneye	✓	✓							
Goldfinch						✓	✓		✓
Goosander	✓	✓		✓					
Grasshopper Warbler									
Great Spotted Woodpecker						✓	✓		
Great Tit						✓	✓		✓
Great Crested Grebe[1]	✓	✓	✓		✓	✓	✓		
Green Woodpecker						✓	✓		

Table II.3.1, continued.

	Scheme								
Species	***Winter WeBS***	***Summer WeBS***	***CBC***		***WBS***	***BBS***			***CES***
			Still	Linear		Still	Linear	Damp	
Greenfinch						✓	✓		✓
Grey Heron						✓	✓		
Grey Partridge						✓	✓		
Grey Wagtail				✓	✓		✓		
Greylag Goose*	✓	✓							
House Martin							✓		
House Sparrow						✓	✓		
Jackdaw						✓	✓		
Jay						✓	✓		
Kestrel							✓		
Kingfisher				✓	✓				
Lapwing					✓	✓	✓	✓	
Lesser Whitethroat							✓		✓
Linnet						✓	✓	✓	✓
Little Grebe[1]	✓	✓	✓		✓	✓	✓		
Long-tailed Tit						✓	✓		✓
Magpie							✓	✓	
Mallard	✓	✓	✓	✓	✓	✓	✓	✓	
Meadow Pipit						✓	✓	✓	
Mistle Thrush						✓	✓		
Moorhen		✓	✓	✓	✓	✓	✓		
Mute Swan	✓	✓	✓	✓	✓	✓	✓		
Nuthatch							✓		
Oystercatcher					✓		✓		
Pheasant						✓	✓	✓	
Pied Wagtail			✓	✓	✓	✓	✓	✓	
Pink-footed Goose	✓	✓							
Pintail	✓								
Pochard	✓	✓							
Red-breasted Merganser	✓	✓			✓				
Red-legged Partridge						✓	✓		
Red Grouse								✓	
Redpoll									✓
Redshank					✓				
Redstart							✓		
Reed Bunting			✓	✓	✓	✓	✓	✓	✓
Reed Warbler			✓	✓	✓		✓		✓
Ringed Plover					✓				
Robin						✓	✓	✓	✓
Rook						✓	✓		

Table II.3.1, continued.

	Scheme								
Species	***Winter WeBS***	***Summer WeBS***	***CBC***		***WBS***	***BBS***			***CES***
			Still	Linear		Still	Linear	Damp	
Ruddy Duck	✓	✓							
Sand Martin					✓				
Sedge Warbler			✓	✓	✓	✓		✓	✓
Shelduck	✓	✓	✓		✓				
Shoveler	✓	✓							
Skylark						✓	✓	✓	
Snipe					✓		✓	✓	
Song Thrush						✓		✓	✓
Spotted Flycatcher							✓		✓
Starling						✓	✓	✓	
Stock Dove						✓	✓		
Swallow						✓	✓		
Teal	✓	✓							
Treecreeper							✓		✓
Tufted Duck	✓	✓	✓		✓	✓	✓		
Water Rail		✓							
Wheatear							✓	✓	
Whitethroat					✓	✓	✓	✓	✓
Whooper Swan	✓								
Wigeon	✓	✓							
Willow Tit									✓
Willow Warbler						✓	✓	✓	✓
Wood Pigeon						✓	✓	✓	
Wren						✓	✓	✓	✓
Yellow Wagtail					✓		✓		
Yellowhammer						✓	✓		✓

Footnotes: 1. For Cormorant, Coot, Great Crested Grebe and Little Grebe, winter WeBS count data were only available from 1986, 1982, 1982 and 1985 onwards, respectively.
2. All geese were regarded as introduced species in respect of summer survey data, but only Canada Geese were also regarded as introduced in winter.

4. ANALYSES

A range of statistical approaches have been applied to bird census data to reveal population changes, many developed with and commonly applied to CBC data (e.g. Mountford 1985, Fewster *et al.* 2000; reviews in ter Braak *et al*. 1994, Thomas 1996). The purpose of the current project was not to evaluate different methods for index generation, so we restricted ourselves to using established, standard methods. The log-linear Poisson regression method (a form of Generalized Linear Model: McCullagh & Nelder 1989), in which bird counts (on a logarithmic scale) are modelled as function of categorical site and year effects with a Poisson error distribution, has gained popularity as the standard method for the analysis of summer monitoring data (ter Braak *et al*. 1994, Thomas 1996). Other techniques such as the Mountford method (Mountford 1985) and route regression (Sauer & Geissler 1990) are based on the same basic model and can sometimes be regarded as special cases of the same general approach. Similar log-linear Poisson regression models have been applied to data from the BBS (Field & Gregory 1999) and CES (Peach *et al*. 1998) schemes. Winter WeBS analyses are routinely conducted using the Underhill method (Underhill & Prŷs-Jones 1994): recent comparative work has shown that indices from this method differ negligibly from those produced by a log-linear Poisson regression model incorporating site, year and month effects (Atkinson *et al*. 2000).

We therefore based all our analyses around log-linear Poisson regression. Models for each scheme and species were fitted using the GENMOD procedure of SAS (SAS Institute, Inc. 1996). Species' indices were discarded if the GENMOD output showed indices and variances varying by several orders of magnitude or repeatedly taking values of zero or one (respectively), showing that the indices were unreliable. Scheme-specific variations around the general analytical protocol are described in Table II.4.1.

Table II.4.1. Details of the Poisson regression models used for each survey scheme.

Scheme	***Analytical Protocol***
CBC	All species for which models ran without errors. Site × year model.
WBS	All species regularly indexed by the scheme. Site × year model.
BBS[1]	All species counted in 30 or more squares per year (for each habitat selection). Birds counted in flight excluded because not closely associated with the habitat. Maximum counts across the two annual visits to a square used. Counts weighted both by the inverse of the proportion of appropriate wetland habitat in the region[2] in which the survey square was found which was surveyed and by the number of transects contributing to the total count from a square. Square × year model.
CES	All regularly indexed species included. Site × year model.
Winter WeBS	All regularly indexed species included. Count sub-sites amalgamated as in routine monitoring. Counts used only from species-specific ranges of months employed in routine monitoring (Cranswick *et al*. 1999). Site × month × year model (site × year only when only one month used).
Summer WeBS	No amalgamation of count sites/sub-sites because counts of different sub-sites (e.g. individual gravel pits within complexes) less likely to feature same birds on different counts in summer. Maximum counts over April-June used (by analogy with BBS/CBC and to avoid passage birds and moult flocks).

Footnotes: 1. The standard, national BBS indexing protocol takes 50 squares per species per year as a minimum requirement; we have used 30 because a smaller number should represent adequately habitat subsets, such as types of wetland. The standard index calculation also includes birds in flight and counts are weighted by the inverse of the proportion of the total number of 1×1km squares in a region[2] which is included in the survey. Our weighting procedure incorporated both an adaptation of the latter appropriate for habitat-specific indexing and consideration of the number of transect sections contributing to the count for a given square such that squares with more of the habitat of interest had more influence on our results.
2. Regions here are BTO membership regions, which are determined by human (i.e. potential surveyor) density and of which there are 127 in Britain and Northern Ireland.

5. COMBINING SPECIES INDICES

We have followed the approach of the headline indicators project to combine the population indices of individual species into multi-species indicators on a scheme-specific basis. Thus, all indices were recalculated as ratios of their values in each year to their values in the first year for which data were available, and the geometric mean index value across species for each year then formed the multi-species index. Where indices for a given species were not available from the start of the optimal time period used (as occurred for four species for WeBS data: Table II.3.1), the index values were standardized with respect to the value for the first year in the time series available for the species (instead of being standardised with respect to one). No weighting was applied in the process of averaging across species.

6. COMBINING SCHEMES

In the headline indicators project, a single source of data for each species was chosen as "best" and that source alone was then used for all the derived indicators involving that species. Both because we were unable clearly to identify a single, "best" scheme for many species (each scheme is biased differently with respect to habitat), and because choosing one scheme would mean discarding significant amounts of useful data, we have used all the available, relevant information on each species. In addition to the scheme- and habitat-specific indicators constructed by simply taking mean index values across species, we combined the data available on particular habitats that were contributed by different schemes. This meant, first, combining the data from several different schemes for single species and, second, combining these species-specific, multi-scheme indices across species.

Indices from different schemes (and their variances: see below) can be combined on a species-specific basis exactly as we have combined species-specific indices within schemes. We have used this approach, but with the modification that the annual indices to be averaged were weighted by the inverse of their variances using a standard formula (Sokal & Rohlf 1995). In this way, the results from the scheme providing the most precise abundance indices contributed most heavily to each habitat-specific indicator for a given species.

Following the headline indicator approach and the scheme-specific indicators described above, we then combined species-specific indices without applying any weighting factors, regardless of the data sources used for each species.

Figure II.7.1. Divisions of England and Wales by region: Regional Development Agencies and water companies. Adapted from maps taken from the DETR and Water UK web sites (http://www.local-regions.detr.gov.uk/rda/map/map.gif and http://www.water.org.uk/companies/watersewage.html).

7. REGIONAL INDICATORS

We have attempted to produce regional multi-species indicators for England, Scotland and Wales (too few data were available for Northern Ireland: see Section II.2) and for two, representative, English Regional Development Agency areas: the North-East (approximately equivalent to the Northumbrian Water region) and the South-East (Figure II.7.1). The former RDA area is characterized by comparatively poor coverage by BTO data, the latter one by comparatively good coverage. The North-East RDA region includes the area covered by the (former) counties of Northumberland, Durham, Tyne & Wear and Cleveland and the South-East RDA region was taken to comprise Hampshire, the Isle of Wight, West Sussex, East Sussex, Kent, Surrey, Buckinghamshire, Berkshire and Oxfordshire. (Technically, RDA regions are made up of a mixture of counties and unitary authorities, but the data sets we used were not classified according to these boundaries.) The North-East RDA region closely matches the area also covered by Northumbrian Water, while the South-East one includes most of the areas covered by each of Thames Water and Southern Water

The regional indices were calculated using the same methods as were used for the national ones, but with subsets of the data available. Time and the available sample sizes (in terms of survey sites from which we had count data) proved to be limiting, so we concentrated on two of the better survey data sets and produced indices for all still waters, all man-made still waters, pits only and reservoirs only from winter and summer WeBS counts for the three countries and two RDA regions listed above. This provided an indication of the likely usefulness of regional sub-divisions in indicator calculations.

8. STATISTICAL CONFIDENCE

No attempt was made to assess the precision of the indicator values generated in the headline indicators project (Gregory *et al.* 1999). The measurement of precision is, however, essential if real changes in an indicator over time are to be distinguished from fluctuations due to random chance. It is not merely a statistical nicety but reflects the basic reason why statistics are a necessary part of modern-day ecology and environmental science. The importance of appropriate and rigorous statistical testing is especially great in respect of analyses that are intended to play pivotal roles in the determination of national conservation policies and the monitoring of their success.

Deriving confidence intervals through a randomisation approach (bootstrapping by site) has become the standard for estimating the precision of individual species' population indices (e.g. Buckland *et al.* 1992, Siriwardena *et al.* 1998). However, this method would be difficult to adapt to multi-species indices because different sets of sites are used (even within the same scheme) to index each species' population and because different schemes could not be combined easily. It would also be prohibitively time-consuming for the complex indicator indices produced here because generating single multi-species indices can take up to 4 hours' computing time and several hundred such indices would be necessary to generate useful confidence intervals. Such methods may become more practicable in the future and have recently been applied to a version of the headline indicator that avoids some of the latter's methodological flaws (Freeman *et al.* 2001).

An alternative is to take the analytical estimates of the variances of each annual population index value for each species that are provided by packages such as SAS and to combine them using standard formulae (see Freund & Walpole 1987, p.150 & 167) to estimate the variance of a multi-species indicator. This approach also allows variances to be combined across survey schemes. The method is less sound statistically than bootstrapping, because of the inherent assumption that the distribution of count data is truly Poisson, when they are actually often (if not always) overdispersed. We have used this analytical approach to estimate annual 95% confidence intervals for each of our multi-species indices. By including an overdispersion correction (calculated from the Pearson goodness-of-fit χ^2 statistic and the number of parameters (site and year effects) in each model: SAS Institute, Inc. 1996) in the species-specific models, we accounted for the effect on the variances of any overdispersion, although not as effectively as bootstrapping would allow.

9. COVERAGE OF WATER BODIES OWNED BY THE WATER INDUSTRY

The summer and winter WeBS data sets are the only surveys considered here which are based on a sampling unit of complete still waters. These data sets are therefore the best suited for extension to a wider range of water industry-owned sites, ideally to all such sites, through additional sampling. To reveal the extent to which the existing data sets cover water industry-owned still waters, we contacted all of Britain's water companies (Figure II.7.1) to request information on the locations (grid references) of the inland still waters they own and matched the lists we were sent to the locations of WeBS count sites. Since many WeBS sites are large and "central" grid references are open to interpretation and subject to observer judgement, we matched WeBS site and reservoir locations by 10×10 and 1×1km squares. Thus, we asked what proportion of 10×10 and 1×1km squares which incorporate reservoirs in each water company region also include a WeBS count site. This approach avoids the subjectivity that would be inherent in a comparison of site names.

III. RESULTS

1. SCHEME- AND HABITAT-SPECIFIC INDICATORS

Sample sizes and data quality allowed habitat-specific multi-species indicators to be generated from the various individual schemes as summarized in Table III.3.1. For comparison, the headline all-habitats indicator was based on 172 species and the headline farmland indicator on 20 species (note, however, that very poor monitoring data such as extrapolated two-point regressions through atlas presence-absence data were used for some of the species in each of these totals).

Plots of the indices for each scheme and habitat are shown in Figures III.1.1-III.1.23. The confidence intervals around all of the indices calculated are narrow enough to suggest that reasonably small changes in absolute index values can reasonably be expected to be identified as statistically significant. Further, fluctuations, which were clearly significant at less than the 5% level (95% confidence intervals not or only slightly overlapping) have occurred within the time series of many of the indicators. However, the extent to which the detectability of any changes is biologically relevant will vary from indicator to indicator: it depends on the nature and number of species contributing data to each one. In other words, statistical significance does not equate to biological significance. Similarly, a lack of statistical significance does not necessarily imply a lack of biological significance. These issues are considered further in the Discussion.

Table III.1.1. Numbers of species contributing to scheme- and habitat-specific indicator indices. Number of species includes both native and introduced species, with the number of introduced species shown in parentheses; "-" indicates that no species met the data quality and/or quantity criteria for inclusion in the index.

Scheme	*Habitat(s)*	*Number of species*
CBC	All still waters	12 (1)
	All man-made still waters*	8 (0)
	Linear waterways	12 (0)
WeBS	All still waters	24 (2)
	All man-made still waters	23 (2)
	Reservoirs	23 (2)
	Gravel/sand pits	18 (1)
Summer WeBS	All still waters	24 (6)
	All man-made still waters	24 (6)
	Reservoirs	23 (6)
	Gravel/sand pits	23 (6)
WBS	Linear waterways	28 (1)
BBS	All still waters	57 (3)
	All damp sites	32 (1)
	All linear waterways	75 (4)
	Large linear waterways	40 (1)
	Small linear waterways	57 (2)
	Water meadows	7 (0)
	Reedbeds	-
	Wet heaths	5 (0)
	Bogs	2 (0)
CES	All damp sites	27 (0)
	Wet scrub	27 (0)
	Reedbeds	23 (0)

*Plot numbers were too small to allow separate CBC indices for pits and reservoirs to be calculated.

Figures III.1.1-III.1.23. National, habitat- and scheme-specific indicator index time series for the maximum periods available. The dashed lines on each plot show 95% confidence intervals estimated by combining analytical estimates of the variances of each contributing species' abundance indices.

14 figure pages

2. HABITAT-SPECIFIC MULTI-SCHEME INDICATORS

Data from multiple schemes were available for the following habitats: all still waters in summer (Summer WeBS, CBC, BBS), reservoirs and pits in summer (Summer WeBS, CBC), all linear waterways (WBS, CBC, BBS) and all damp sites (CES, BBS). Plots of the variation over time in multi-scheme indices derived from data from these four habitats are shown in Figures III.2.1-III.2.4.

The multi-scheme indicator indices show levels of statistical confidence similar to those found for the single scheme indicator. In some cases (e.g. Figures III.2.1 and III.2.3), the increased precision introduced by the addition of one or more sources of data mid-way through the time series can clearly be seen. However, the caveats about the interpretation of the confidence intervals with respect to the biological meaning of the indicators that are described above (Section III.1) also apply here.

Figures III.2.1-III.2.4. National habitat-specific indicator index time series derived by combining, first, scheme-specific indices by species before combining the results across species. The dashed lines on each plot show 95% confidence intervals estimated by combining analytical estimates of the variances of each contributing scheme's and species' abundance indices.

3. REGIONAL INDICATORS

The multi-species indicators from winter and summer WeBS data for all still waters, reservoirs and gravel/sand pits, reservoirs only and pits only for each of England, Scotland, Wales and the North-East and South-East regions of England are shown in Figures III.3.1-III.3.17 (winter) and in Figures III.3.18-III.3.36 (summer). The number of species contributing to each regional indicator is shown in Table III.3.1. No species were indexed adequately by winter WeBS data for Scottish, Welsh or North-East English gravel/sand pits alone or by summer WeBS data for North-East English pits. In addition, the summer data for Scottish and Welsh pits were insufficient to allow indices for 1994 (the first year used nationally) to be calculated, suggesting that the data available approach the threshold below which the generation of an indicator would not be feasible. A similar potential limit is likely to apply to the generation of regional indicators based on specific habitats other than wetlands.

The confidence intervals around the indicator indices, as shown in Figures III.1.1-III.1.36, were comparable to those generated for the national indices, suggesting that these sub-divisions can also provide statistically viable indicators, despite the smaller numbers of species included. However, it is important to note that the biological value of the indicators will have been reduced by the use of a smaller range of species. This issue is discussed in Section IV.1.

Figures III.3.1-III.3.36. Region-and country-specific indicator index time series for the maximum periods available. The dashed lines on each plot show 95% confidence intervals estimated by combining analytical estimates of the variances of each contributing species' abundance indices.

29 figure pages

Table III.3.1. Numbers of species contributing to regional habitat-specific indicator indices derived from winter and summer WeBS data. Number of species includes both native and introduced species, with the number of introduced species shown in parentheses; "-" indicates that no species met the data quality and/or quantity criteria for inclusion in the index.

Scheme	***Region***	***Habitat(s)***	***No. of species***	***% of no. in national index***
Winter WeBS	England	All still waters	24 (2)	100
		All man-made still waters	22 (2)	96
		Reservoirs	21 (2)	91
		Gravel/sand pits	19 (1)	100
	Wales	All still waters	17 (0)	71
		All man-made still waters	12 (0)	50
		Reservoirs	12 (0)	50
		Gravel/sand pits	-	-
	Scotland	All still waters	19 (0)	79
		All man-made still waters	16 (0)	67
		Reservoirs	16 (0)	67
		Gravel/sand pits	-	-
	North-East England	All still waters	16 (0)	67
		All man-made still waters	11 (0)	46
		Reservoirs	11 (0)	46
		Gravel/sand pits	-	-
	South-East England	All still waters	18 (1)	75
		All man-made still waters	18 (1)	78
		Reservoirs	16 (0)	70
		Gravel/sand pits	16 (1)	89
Summer WeBS	England	All still waters	24 (6)	100
		All man-made still waters	24 (6)	100
		Reservoirs	22 (5)	92
		Gravel/sand pits	23 (6)	100
	Wales	All still waters	18 (3)	75
		All man-made still waters	10 (1)	42
		Reservoirs	10 (1)	43
		Gravel/sand pits	3 (0)	13
	Scotland	All still waters	22 (4)	92
		All man-made still waters	14 (1)	58
		Reservoirs	13 (0)	57
		Gravel/sand pits	5 (0)	22
	North-East England	All still waters	19 (3)	83
		All man-made still waters	14 (2)	58
		Reservoirs	14 (2)	61
		Gravel/sand pits	-	-
	South-East England	All still waters	19 (4)	79
		All man-made still waters	19 (4)	79
		Reservoirs	15 (2)	65
		Gravel/sand pits	19 (4)	83

4. COVERAGE OF WATER BODIES OWNED BY THE WATER INDUSTRY

The results of the comparison of the locations of water company reservoirs and WeBS count sites are summarized in Table III.4.1. They suggest WeBS coverage is generally good: the comparatively poor match for 1×1km squares probably partly reflects grid reference differences between count points and the centres of reservoirs. (The difference between water companies in the percentage match at this scale could also reveal more about the exact system used by different companies to record reservoir grid references than it does about regional differences in WeBS coverage.) It is also notable that, if a reservoir is not counted but other counts are made in the same 10×10km square, the counts are likely to sample the populations of waterfowl that use the reservoir. This is especially likely to be true in winter.

Table III.4.1. Summary of the coverage of water-industry owned reservoirs by WeBS counts. Data are presented by 10×10km square and 1×1km square: reservoirs were considered to be covered by WeBS if a WeBS count site fell in the same square at the appropriate scale. Note that no data were obtained from water companies not listed in the table: a significant proportion of the reservoirs owned by these companies are probably covered by WeBS, improving the coverage figures given below for the whole of the UK ("overall"). Note that two large water companies, Wessex and South-West Water, were not contacted.

Water Company	***10×10km square level***			***1×1km square level***		
	No. of squares with reservoirs	***No. also with WeBS sites***	***%age covered by WeBS***	***No. of squares with reservoirs***	***No. also with WeBS sites***	***%age covered by WeBS***
Anglian	14	11	78.6	17	3	17.7
Northumbrian	13	13	100	23	9	39.1
North West	55	54	98.2	155	110	71.0
Severn Trent	25	21	84.0	30	14	49.7
Southern	4	4	100	4	0	0
Thames	89	83	93.3	244	22	9.0
Welsh	48	39	81.3	89	32	36.0
Yorkshire	40	34	85.0	111	53	47.8
OVERALL	288	259	89.9	673	243	36.1

IV. DISCUSSION

1. WHICH REGIONS AND HABITATS DO THE INDICATORS COVER ADEQUATELY?

We have been able to construct reasonably statistically precise multi-species indicators in all the scheme-, habitat- and region-specific analyses for which two or more species met our data quality and quantity criteria. The only cases where indices could not be calculated were BBS reedbeds (although note that an index was available for this habitat from the CES), gravel/sand pits for Wales, Scotland and North-East England (Winter WeBS) and gravel/sand pits for North-East England (Summer WeBS) (Tables III.1.1 & III.3.1). Indicators could not be produced for some fine habitat sub-divisions, such as gravel pits or reservoirs alone for CBC data, or for data-poor areas such as Wales or the North-East England region. This suggests that additional data collection or the amalgamation of regions would be necessary for indicators equivalent to those generated for data-rich areas such as South-East England to be produced.

Notwithstanding the problems with some regions and habitats, the results suggest that the use of the indicators generated for the majority of habitats is broadly feasible, *statistically*. However, this raises a critical point, namely the extent to which the indicators represent valid indicators *biologically*, i.e. what constitutes "adequate coverage". This will depend on the features of the appropriate wetland environment that are of interest in terms of their sustainability and on the relationships of the abundances of the species included in the indicator to these features; it has no necessary relationship with statistical precision. Requirements for the identification and characterization of such relationships are discussed in Section IV.2.

It is also important to consider the number of species that are included in each index: statistical precision may be high, but if only a small proportion of the species that use a given habitat in given region is monitored adequately, important bird population or habitat changes might be missed. If this proportion is small because only a few species are common in the habitat/region concerned (with other species using the habitat but being scarce and therefore not well surveyed), there may be no problem. However, if several species are missed because they are difficult to monitor (e.g. cryptic or nocturnal birds), additional surveying might be required to ensure an adequate coverage of the bird community.

Returning to statistical issues, an obvious question is whether an indicator's sample size provides sufficient power to allow a change of a given size (which has been determined to be of interest, *a priori*) to be detected. In the planning stages of this project, this question was phrased in terms of the number of survey sites contributing to an indicator. Unfortunately, the nature of the multi-species and multi-scheme monitoring data is such that sample sizes cannot be expressed so simply, for several reasons. First, even when only a single scheme is involved, the indices for each species are calculated from a different set of sites, each site holding a different set of species (in varying proportions): adding further sites would therefore have no easily predictable effect and the number of contributing sites *per se* is not a particularly informative statistic. More fundamentally, *sample size* is, in the context of multi-species indices, some combination of the number of sites and the number of species. While the biological meaning of an indicator depends critically on the ecological characteristics of the species used in an indicator, even without ecological knowledge, a clear general rule will be that more species and more sites will give more ecological, regional and habitat coverage. The form of this relationship is difficult to predict, however, because neither species nor sites differ in regular ways, such that only rarely could two sets of an equal number of sites or species ever be said with certainty to provide equivalent monitoring information. As an illustration, an indicator based on two ecologically similar species such as Tufted Duck *Aythya fuligula* and Pochard *A. ferina* would probably detect certain changes more readily than one based on a more disparate pair of species such as Mallard *Anas platyrynchos* and Tufted Duck, but the latter might have the potential to detect a wider range of changes.

One general characteristic that varies between species illustrates the difficulties inherent in quantifying sample size in terms of species. First, habitat generalists make up much of the sample for some habitat divisions in analyses such as those of BBS data (Table II.3.1): the abundance of such species will be less closely linked to environmental conditions and will therefore provide a less effective, or at least a less sensitive, indicator than the abundance of specialists on the habitats of interest. A sample size of, say, ten generalists can therefore not be regarded as equivalent to one of ten specialists. This issue has particular relevance in comparatively data-poor habitats and regions: our indicators in such cases typically chiefly comprised indices for common generalist species such as Skylark *Alauda arvensis* and Chaffinch *Fringilla coelebs*. The small species numbers available for habitats such as wet heaths and water meadows (Table II.3.1) might even, therefore, tend to overestimate the power of the multi-species indices as indicators compared to an (hypothetical) indicator based on specialists.

Finally, a further fundamental point to which we have alluded above is that the size of change required to be detected must be decided *a priori* before sample size questions are relevant. If a multi-species index is merely to be used to *indicate* average population trends, this is an arbitrary policy decision: does an average decline of (for example) 5, 10, 25 or 50% “matter”? If an index is to indicate the effects of particular environmental influences, the value set as an important change should be supported by data on quantitative responses of the indicator to known environmental changes. In other words, we need to know the meaning of a change of x% in an indicator before it is meaningful to ask whether such a change is detectable. This means either waiting for changes to occur and watching what the indicator does or investigating how the indicator responded to known, historical events. For the latter, spatially referenced time series of data contemporary with BTO bird surveys that describe important changes in water management, which have occurred in the past (such as, perhaps, abstraction rates, water quality, amounts of recreational use and canal dredging frequencies) would be required. Without calibration work such as this, we cannot define the sample sizes required for an indicator to be effective: the qualitative and quantitative meanings of changes in an indicator must be determined first. Then, a decision about the size and direction of change that is of interest can be made, leading to consideration of whether the data available permit such changes to be detected. The issue of whether multi-species indices are better regarded as indicators of environmental factors or merely of average bird population trends is discussed further below.

It may be of interest to compare the indicators we have produced with the headline wetland indicator time series, and this is shown in Figure IV.2.1 (taken from Gregory *et al.* 1999). In common with most of our long-term indicators, this index shows a continuing smooth increase over time, reflecting the overriding pattern amongst UK wetland bird populations (see, e.g., Cranswick *et al.* 1999). Whether or not our indicators based on the broader habitat divisions represent better indicators of the health of UK wetlands than the headline index cannot be determined at this stage: the relationship each indicator has with changes in the environment would have to be determined first.

Figure IV.2.1. The trend shown by the headline indicator for all lowland wetland species. Reproduced from Gregory *et al*. (1999). No confidence intervals are available for this index.

One figure page

2. WATER INDICATORS AND SUSTAINABILITY

We have shown that multi-species indices can be calculated for most wetland habitat-region combinations and that these indices are reasonably statistically precise. So, to what extent can these indices be regarded as indicating the sustainability of management by the water industry? The answer to this question was alluded to in the discussion of calibration work above: the indices we present show average population trends tied as closely as possible to wetlands by habitat and/or species selections, but we have no information on how management might affect these average trends. Thus, a multi-species index might show long-term stability reflecting the stable populations of a range of more generalist (or merely less severely affected) species whilst management practices are causing declines in one or two specialists. In such a situation, the overall pattern would suggest that management is sustainable but this would certainly not be true for the specialists concerned, whose population declines are masked in the multi-species index by those of the other species. Wetland species with specialist ecologies or breeding habitats that could be affected in this way include Little Ringed Plover *Charadrius dubius*, Kingfisher *Alcedo atthis* and Dipper *Cinclus cinclus*.

To identify how well changes in a multi-species index reflect the environmental effects of water industry management practices (and therefore the sustainability of the latter), the performance of the putative indicator needs to be assessed in the face of known environmental changes. This could potentially be done in two ways: through correlation of historical indicator time series with historical environmental data or through monitoring future changes in the environment and the indicator in parallel. In practice, the important issue is probably whether the rather blunt instrument of a multi-species average population trend would respond to given changes of conservation concern so that a problem would be detected using such an index (i.e. the risk that a Type II error would be made), rather than whether the average trend will show significant changes that are unrelated to water industry management.

A major impact on species abundance in many habitats can be habitat succession. There is an in-built control for these changes in national, randomised schemes such as the BBS, but site-based and habitat-specific monitoring can be affected. This has particular relevance to the monitoring of the sustainability of water industry management because reservoirs and gravel pits tend to mature gradually as wildlife habitats over decades. Numbers of many species would therefore be expected to show clear long-term trends (perhaps mostly increases) as peripheral vegetation, for example, develops. These patterns would have to be accounted for in the interpretation of any average population trend, in order to reveal the role that water industry management might have played in determining the trend.

Notwithstanding the above discussion, a stable multi-species average population trend (note that an increasing index is no more *sustainable* than a declining one) could be chosen as the criterion indicating sustainability. This would imply that management had been sustainable if an average species had not changed in abundance. Careful inspection of confidence intervals (i.e. checking that they are not too wide) would show whether the average value masked many non-sustainable increases or declines, but large declines or increases of individual species could still be occurring and going unnoticed (the possible number will rise with the number of species contributing to the index), as described above. The key point here is that the *average* of the trends of set of species would have to be accepted as being the measure of interest, thereby ignoring any species-specific effects that are too small to affect that average.

3. ADDITIONS, ALTERATIONS AND ADAPTATIONS

There are several ways in which changes to the methods we have used might produce better indicators and these should be considered in any further work on indicator generation and in their application. First, the inevitable differences between species in the extent to which indices of their abundance reflect environmental conditions suggests that species-specific weights should be applied when indices are combined. Such weights would best be determined by investigations of each species' ecology and the influences of the key features of the environment on abundance. Species with shorter generation times and habitat specialists would be expected to respond more quickly and more unequivocally to environmental change and would therefore merit larger weights. Likewise, species that are resident in the UK will potentially provide information on the UK environment throughout the year while migrants will be affected by many overseas factors in addition to factors of direct interest to the UK: residents might therefore merit larger weights than residents. Once the environmental changes that the indicator is required to track have been determined, it should be possible to find the information necessary in the literature to assign weights on these bases. An alternative to the application of quantitative weights is to select species by habitat specialization or likely speed of demographic response, or to choose a suite of species each of which has a known response to the environmental influences in question.

Another context in which weights might be applied is with respect to the statistical reliability of the data, perhaps using the inverse of the variance of an index value (as used here to combine indices by species: see Section II.5). This would prevent the potential problem of numerically large but statistically uncertain changes in index values having important effects on an indicator (as well as the converse) which could be misleading. This problem is perhaps potentially more serious within the approach used for the headline indicator (Gregory *et al.* 1999) because of the wider range of data quality found there.

A basic presumption of the entire multi-species indicator concept is that larger numbers of each and every species are desirable and that declines are always bad. This is, however, unlikely to be true even in the monitoring of biodiversity. It may be that the desirable direction of population change for some species tends to be downwards. As an illustration, eutrophication is usually considered to be an undesirable process in rivers and still waters, yet it is associated with a very high abundance and biomass of taxa such as green algae and chironomid flies. Among birds, species that are common in (more broadly) poor quality habitats, such as Wood Pigeons *Columba palumbus* on farmland, might reasonably be given negative index values contributing to an indicator of environmental quality. In other cases, population trends might be known to be unrelated to the UK environment. For example, declines in UK Ringed Plover *Charadrius hiaticula* populations in winter might actually reflect improving conditions in continental Europe, such as could have been induced by changes in climate (M.M. Rehfisch & G.E. Austin, pers. comm.). In such a case, a declining indicator could reflect no change in UK wetlands and/or improvements elsewhere, but might nevertheless be interpreted as showing deteriorating conditions in the UK. Such species might be omitted from a multi-species index to maximize the relevance of the index to the regions and habitats of interest.

One long-running and reasonably strong data set which could be included in future indicators for wetland habitats is that derived from the BTO Heronries census, which has recently been re-analysed using more refined statistical methods than were applied to it previously (Freeman *et al*. 1999). We did not include this data set because it is based on counts of nests in colonies from which the birds surveyed are likely to fly considerable distances to a range of foraging locations, making associations of habitat with the counts difficult. However, this may not be a problem in other monitoring applications, such that the heronries data could be a valuable ingredient in an indicator.

We also omitted all data on introduced species from our indicators, following the headline indicators project (Gregory *et al*. 1999). This omission can be justified on the basis that most conservationists would regard a healthy environment as being characterized by native species and that increasing or stable indicators whose values are driven by the abundances of introduced species could mask declines in the native community and therefore be misleading. However, there is no reason, *a priori*, why native species should be a better indicator of, say, the effects of abstraction, than introduced ones. If introduced species are undesirable, their abundance indices could perhaps be given negative weights in the indicator production process. In general, the decision whether to include a given species (native or introduced) should be based on its relationships with the features of the environment which are of interest. Whatever decision is made with respect to the inclusion of introduced species, an issue that should be addressed is that there are a number of species, especially among wetland birds in the UK, whose populations are or have been substantially augmented or supported by releases (some accidental) of captive-bred birds. The UK populations of species such as Mallard *Anas platyrynchos*, Gadwall *A. strepera*, Grey Partridge *Perdix perdix* and Pheasant *Phasianus colchicus*, for example, might therefore be regarded as introduced just as justifiably as species such as Ruddy Duck *Oxyura jamaicensis* and Canada Goose *Branta canadiensis*. Whether such species are included in any indicator will depend on the aims of the analysis and could well be a species-specific decision.

One procedural change that would improve the interpretability of a multi-species index would be to present it in conjunction with its constituent species-specific trends. This would enable the reader to see, at a glance, how consistent the patterns underlying an overall trend are. The principal difficulty with such an approach would be with respect to presentation. This might be alleviated by presenting a large number of species-specific trends on single graphs (deviations from a common pattern would be of most interest, so confusion between species that have undergone similar trends would be unimportant) or by building links to species graphs into web-based presentations. Front-line publications could still use just the average trend, but any documents discussing or interpreting the pattern it showed should also present the constituent information.

A final potentially beneficial addition to the methods we have used would be to include an explicit consideration of the problem of the loss and gain of sites from monitoring schemes. A general problem with any habitat-specific monitoring is that it, by definition, only provides information about that particular habitat: sites typically cease to be monitored once a habitat, such as a pond, ceases to exist and are almost always not monitored *before* the habitat of interest appears. Schemes that are random with respect to habitat, such as the BBS, do not have this problem (although it is created even in BBS data when habitat-specific subsets of the counts are extracted). The problem is only avoided if the spread of habitat coverage represented by the sites contributing to a scheme remains constant over time: this has essentially been the case, for example, for the farmland CBC (Marchant *et al.* 1990).

Two approaches are available to counter the problem of changes in the habitat composition of the monitoring sample. First, comparison of the abundance indices for habitat specialists from habitat-unbiased schemes with those from habitat-specific monitoring would reveal whether important changes had been missed because sites from which the species had been lost would not disappear from the former data set. (Sites from which a species has disappeared would subsequently contribute zero counts to the indexing process.) A second option would be the monitoring, independently of that of bird populations, of the existence of sites. Information from surveyors as to why they start or stop covering particular sites could contribute to this process, as could periodic return visits to sites that are no longer monitored to assess whether large-scale habitat changes have occurred.

4. WHAT IS THE VALUE OF MULTI-SPECIES INDICATORS?

An annual, single figure in which changes in bird populations, perhaps indicating environmental conditions, can readily be seen would be a valuable tool with which managers and conservationists could both monitor the impacts of potential influences on the environment (positive and negative) and communicate their findings effectively and as simply as possible to government and the general public. Multi-species indices of abundance were developed for the government's headline indicators project, and were subsequently adopted in the determination of policy, with rather limited discussion and consultation among conservation scientists. Many conservationists and managers not closely associated with the work behind the method used to generate the headline indicator will assume that an index that has achieved such a high profile must have been tested thoroughly and found to be effective. It is important that such users should understand how such indices are constructed and that there are important limitations and caveats that must be incorporated into any interpretation of them. We recommend, therefore, that discussions are held, both within and across the key organizations, to re-consider the issues surrounding multi-species indicators and that outside agencies that are interested in developing indicators at regional or habitat-specific levels do so while viewing the headline project approach as one, imperfect, possibility rather than as a recommended template. A full review of the issues here would be beyond the scope of this report (and would require far wider consultation), but we have already considered some methodological points and the principal conceptual ones are summarized below.

In the ecological literature, the term "indicator" generally refers to a biological measure (usually the abundance of specified organisms or groups of organisms) that can be used as a surrogate for less tractable and more fundamental environmental variation (including human impacts). Examples include bird indicators of mining and grazing impacts (Read *et al.* 2000) and aquatic plant indicators of water quality and human use (Nichols *et al.* 2000): in such cases, potential indicators are tested against real data and, critically, some of them can fail the test, proving to be of little use (see, e.g., Read *et al.* 2000). The process of developing and testing indicators is significant enough, in itself, to support the launch, in 2001, of a journal dedicated to it (Elsevier's *Ecological Indicators*). In one example, the development of indicators of land quality (considering human impacts on the environment in combination with monitoring of the economic performance of managed ecosystems) from conception, through planning, to validation has been subjected to peer review and conducted in the public domain (Dumanski 2000, Dumanski & Pieri 2000, Hurni 2000, Steiner *et al.* 2000). The foregoing shows that the development of meaningful environmental indicators is (and should be) a carefully conducted and often complex process. A common argument against sophistication in indicator development is that simplicity, i.e. ease of understanding of the method and its application by the public, is important. This is true, but should not override the fundamental requirement that is the development of a good index that supplies the information required as efficiently as possible. Simplicity would be a valid criterion to use to choose between equally meaningful candidate indices, but there can be little value in an index that is simple mathematically but uninterpretable ecologically.

In the context of the work described in this report, it is important to remember that without the demonstration of relationships with particular environmental features, a multi-species indicator *merely shows an average population trend*. The term "indicator" might therefore be considered to be a misnomer unless the index concerned is shown to *indicate* something other than its own value. This would suggest that considerable testing of the behaviour of any multi-species index under changing environmental conditions would be necessary before changes in the index could be interpreted with respect to conservation priorities, such as whether management is sustainable. However, an index might be required only to be an indicator of the general health of bird populations: such a goal might be served well, in principle, by an average population trend, but the properties of such average trends need to be understood and accepted before they are adopted as monitoring tools.

First, a simple average trend implicitly assumes that all species' population changes are equivalent, so there is no differentiation between species with different conservation values, likely speeds of response to environmental change, geographical ranges or ecological (i.e. habitat) specializations. The set of species that contribute to an index will be influenced not only by whether species are common in the habitat or area of interest but also by the ease with which they can be surveyed by existing schemes. This means that the index will not be an all-inclusive summary of appropriate bird populations.

Second, combining species blindly will produce a *simple* index, but the meaning of the index in terms of all bird populations will be vague and uncertain, rather than a general summary. Many of the individual species indices, such as those from the farmland CBC or the BBS, have known or controlled habitat or regional representation (or biases) and, in some cases, established demographic equations underlying the long-term trends. This makes these trends more interpretable. Averaging such trends across species will produce less interpretable patterns: in a real sense, less than the sum of their parts.

Third, in most cases, the simple averaging of disparate species will produce something that responds slowly to environmental change, because the earliest responses will be by the most vulnerable species and such changes will be swamped by the others contributing to the index until many of the latter also respond similarly. There will always be some lag been environmental cause and species abundance effect, but it will usually be desirable for this lag to be as short as possible. The averaging of species' trends will always tend to "iron out" fluctuations to produce rather flat lines (as shown by the long-term trends produced for this report). In the context of sustainability, this means that indices will be very conservative, always tending to suggest that management has been "sustainable". An averaged index will only achieve an "early warning" function if combining species reveals trends not apparent in individual species' trends; it is hard to imagine how a simple average could ever do this.

Fourth, an agency adopting an averaged index has an implicit policy that all population changes are equivalent. To illustrate what this means, a situation where nineteen of twenty species in an index had stable populations but the twentieth had halved in abundance could (depending on statistical significance) represent an issue of conservation concern or show a lack of sustainability. However, it would be concluded that there were no conservation problem or that management was sustainable if one of the other nineteen species' populations had doubled over the same period, rather than remaining stable. In the farmland context, it is plausible that the declining species might be Yellowhammer *Emberiza citrinella* and the increasing one Woodpigeon *Columba palumbus*: should these population changes be considered equivalent?

Fifth, it might be tempting to view averaged trends based on several species' abundance indices as measures of biodiversity. It is important to note that they would be very poor for this purpose. For example, stability in nine of ten species and a 50% increase in the tenth would result in the same increase in a geometric mean population trend as a 4.14% increase in each of the ten species. By most definitions, the latter would represent more of an increase in biodiversity (and would be more desirable in terms of conservation). As an aside, it would be quite possible to develop long-term indicators of biodiversity *per se* from BTO survey data by calculating standard diversity indices (or modifications thereof) for individual survey plots and combining them regionally or nationally in an appropriate framework. This would form a project in its own right.

Sixth, even if an averaged trend is viewed, formally, as providing information only about species trends and not necessarily about the environment, there is a danger that it will be misinterpreted as a general indicator of environmental conditions (a "magic number"). There is a danger that informal conclusions such as "the indicator is stable, so the environment must be okay" will be reached from published averaged bird population trends. It should be emphasized to agencies adopting averaged trends as indicators and to the end users of these indicators that such conclusions are not necessarily valid.

Finally, we must consider whether multi-species indicators based on bird abundance data have proven to be useful in practice. Of the indices developed in the headline indicator project, the farmland-specific indicator has achieved the highest profile. This index was developed long after widespread declines in farmland bird populations had been identified (see, e.g., O'Connor & Shrubb 1986, Marchant *et al*. 1990, Fuller *et al*. 1995, Siriwardena *et al*. 1998) and after they had been accepted as fact by conservationists, many policy-makers and much of the wider public. However, its inclusion in the UK government's set of "quality of life" indicators resulted both in the problem becoming still more widely recognized and in the government making a policy commitment to reversing the decline in the farmland bird indicator by 2020. Further, the woodland bird indicator has drawn wider attention to the declines in many woodland bird populations that, having been less dramatic than the declines on farmland, had previously received little attention outside a limited group of specialists. These indicators have therefore played a potentially important role in publicizing the results of bird monitoring and in focusing public and government attention on certain key results. However, it is important to remember that they have not yet identified any problem that was not previously known. The successes of the headline indicators therefore represent evidence only that such indices can be useful as tools for increasing publicity and not evidence that they are useful tools for *monitoring* the environment. In other words, the available evidence suggests that simple multi-species indicators can be useful in the dissemination of the results of monitoring but not within the monitoring process itself. In the absence of validation work, it would be wise to use simple multi-species indices only to summarize existing results, the presentation of the results being explicit about their limitations as monitoring tools.

5. FURTHER WORK

The results of this project and the foregoing discussion suggest several important directions for future work. The most valuable piece of work would be a consultation exercise to determine what direction the future development of multi-species bird indicators should take. This work should build on the discussion above and the issues of statistical validity and biological usefulness we have identified. In particular, it should be considered whether a multi-species approach or one derived from data on the abundance or another ecological or demographic feature of a single species is the more useful. Issues to be considered for multi-species indices would include how best to determine the list of species to include and weights for species-specific indices when they are combined into averaged trends. One central issue is whether summaries of bird populations trends *per se* are of interest or whether an index should be *indicative* of features of the environment or its health. If the latter is the case, it is important that appropriate validation studies are conducted so that the correct inferences about the environment can be drawn from changes, or a lack of change, in an "indicator". Such studies would be critical, for example, if stability in a wetland-based index were to be used to indicate sustainability in water industry management: we would need to know under what conditions changes in the index are likely to occur and, perhaps more importantly, how bad conditions would have to be before changes become detectable. Depending on the use to which an indicator would be put, this might only require collating information from the literature and historical data sets, but it is also possible that the collection of new data on the appropriate environment-abundance relationships would be necessary.

If the aim of an indicator is to produce indices that allow the monitoring of general environmental conditions (including the sustainability of management), indicators should be developed using specific relationships with environmental variables and then tested, as described above and in the growing literature on the subject. It is likely that habitat-specific abundance indices for single species, or perhaps for small groups of species, will prove to be the best, most sensitive measures of environmental impacts and therefore of whether management is sustainable.

If the aim of an indicator is just to summarize existing results, alternative approaches to formal indicator generation exist that would avoid the potential problems of misinterpretation with the multi-species, average index approach and may therefore merit further consideration. Simple statistics such as the number of species declining or increasing in one habitat as opposed to another or the mean population change over a set period for species which are specialists on the habitat of interest versus that for generalist species would provide as much reliable information as a multi-species indicator and be more transparent. The significance of such statistics could then be assessed using, for example, χ^2, G- or t-tests. Which statistic and which test is used would best be selected according to the quality of the monitoring data: a more quantitative approach, using the absolute size of population changes, would be more appropriate where monitoring data are good; qualitative comparisons based, say, on numbers of species which have undergone statistically significant population changes would be more appropriate where confidence intervals around population changes are large. This type of approach would, additionally, make the set of species used more obvious to an end user, thus addressing (in part) the issues discussed in Section IV.3 above. An example of the application of a comparison of mean population changes between specialist and generalist species on farmland can be found in Siriwardena *et al.* (1998).

As discussed above, the success of the farmland bird indicator has been to publicize established results. If this is the sole aim of further indicator development, consideration should be given to the possibility that the best course of action is to focus on improving the dissemination of results to government and the wider public. It is difficult to make sensible, specific recommendations without specialist knowledge in the area of publicity, but ideas might include making greater use of the internet and placing a high priority on the publication of results in peer-reviewed and widely read scientific journals, with spin-off articles in the popular and semi-popular press, rather than in research reports with a restricted distribution. The BTO has recently begun to use the Internet to supply information on bird populations to the general public with its latest report on *Breeding Birds in the Wider Countryside* (Baillie *et al*. 2001) and initial responses have been positive and promising for the future.

ACKNOWLEDGEMENTS

This study was funded by Northumbrian Water and we thank Chris Spray for enabling the work to proceed. We are very grateful to The Wildfowl and Wetlands Trust and WeBS Secretariat staff, notably Peter Cranswick, Becky Hughes and Mark Pollitt, as well as Andy Musgrove, Graham Austin and Phil Atkinson of the BTO, for their help in providing and interpreting WeBS data. Dawn Balmer, Richard Bashford, John Marchant, David Noble and Fiona Sanderson (BTO) helped us extract and interpret CES, CBC, BBS and WBS data. Mark O'Connell and Chris Spray provided helpful comments on an earlier draft. Steve Freeman and Stephen Baillie provided valuable statistical advice. David Gibbons and Richard Gregory (RSPB) supplied useful information regarding the headline indicators project. Finally, we thank Stuart Pudney & Chris Spray (Northumbrian Water), Roy Rhodes (North West Water), Phil Morgan (Welsh Water), Andy Brown & Ken Smith (Anglian Water), Miles Foulger (Yorkshire Water), Julie Healy (Severn Trent Water), Tom James (Southern Water) and Vicky Bower (Thames Water) for supplying data on the locations of water industry-owned reservoirs.

REFERENCES

Atkinson, P.W., Austin, G.E., Burton, N.H.K., Musgrove, A.J., Pollitt, M. & Rehfisch, M.M. (2000) *WeBS Alerts 1998/99: Changes in numbers of wintering waterbirds in the United kingdom at national, country and Special Protection Area scales.* BTO Research Report 239, British Trust for Ornithology, Thetford.

Baillie, S.R., Crick, H.Q.P., Balmer, D.E., Bashford, R.I., Beaven, L.P., Freeman, S.N., Marchant, J.H., Noble, D.G., Raven, M.J., Siriwardena, G.M., Thewlis, R., Wernham, C.V. (2001) *Breeding Birds in the Wider Countryside: their conservation status 2000.* BTO Res. Rep. no. 252. British Trust for Ornithology, Thetford, Norfolk. (http://www.bto.org/birdtrends)

Baillie, S.R., Green, R.E., Boddy, M. & Buckland, S.T. (1986) *An evaluation of the Constant Effort Sites scheme.* Unpublished report to the Ringing Committee of the British Trust for Ornithology.

Baillie, S.R., Gregory, R.D. & Siriwardena, G.M. (1997). Farmland bird declines: patterns, processes and prospects. In *BCPC Proceedings No. 69: Weeds 1997*, ed. R.C. Kirkwood, BCPC, Farnham, UK.

Buckland, S.T., Cattanach, K.L. & Anganuzzi, A.A. (1992) Estimating trends in abundance of dolphins associated with tuna in the eastern tropical Pacific Ocean, using sightings data collected on commercial tuna vessels. *Fishery Bulletin*, **90**, 1-12.

Cranswick, P.A., Pollit, M.S., Musgrove, A.J. & Hughes, R.C. (1999) *The Wetland Bird Survey 1997-98: Wildfowl and Wader Counts.* BTO/WWT/RSPB/JNCC, Slimbridge, UK.

Dumanski, J. (2000) Land quality indicators. *Agriculture, Ecosystems and Environment*, **81**, 81.

Dumanski, J. & Pieri, C. (2000) Land quality indicators: research plan. *Agriculture, Ecosystems and Environment*, **81**, 93-102.

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.

Field, R.H. & Gregory, R.D. (1998) *Habitat bias in Actual and Ideal Transect Lines in the Breeding Bird Survey, 1994-1997.* BTO Research Report No. 214. British Trust for Ornithology, Thetford.

Field, R.H. & Gregory, R.D. (1999) *Measuring population changes from the Breeding Bird Survey*. BTO Research Report No. 217. British Trust for Ornithology, Thetford.

Freeman, S.N. (1999*) A comparison of methods for analysing BTO Heronries Census data.* Unpublished report to the BTO.

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

Freund, J.E. & Walpole, R.E. (1987) *Mathematical Statistics, 4th Edition.* Prentice-Hall International, New Jersey, USA.

Gibbons, D.W., Reid, J.B., & Chapman, R.A. (1993). *The New Atlas of Breeding Birds in Britain and Ireland: 1988-1991*. Poyser, London.

Gregory, R.D. & Baillie, S.R. (in press) Survey design and sampling strategies for breeding bird monitoring. *Proceedings of the 13th International Conference of European Bird Census Councils, Estonia 1995.*

Gregory, R.D., Baillie, S.R. & Bashford, R.I. (in press) Monitoring breeding birds in the United Kingdom. *Proceedings of the 13th International Conference of European Bird Census Councils, Estonia 1995.*

Gregory, R.D., Gibbons, D.W. Impey, A.J. & Marchant, J.H. (1999) *Generation of the headline indicator of wild bird populations.* BTO Research Report No. 221. British Trust for Ornithology and Royal Society for the Protection of Birds, Thetford.

Hurni, H. (2000) Assessing sustainable land management (SLM). *Agriculture, Ecosystems and Environment*, **81**, 83-92.

Marchant, J.H., Hudson, R., Carter, S.P. & Whittington, P. (1990) *Population trends in British breeding birds*. British Trust for Ornithology, Tring, Hertfordshire.

McCullagh, P. & Nelder, J.A. (1989) *Generalized Linear Models, 2nd Edition*. Chapman & Hall, London.

Mendoza, G.A. & Prabhu, R. (2000) Development of a methodology for selecting criteria and indicators of sustainable forest management: a case study on participatory assessment. *Environmental Management*, **26**, 659-673.

Mountford, M.D. (1985) An index of population change with application to the Common Birds Census. *Statistics in Ornithology* (eds B.J.T. Morgan & P.M. North), pp 121-132. Springer-Verlag, Berlin.

Nichols, S., Weber, S. & Shaw, B. (2000) A proposed aquatic plant community biotic index for Wisconsin lakes. *Environmental Management*, **26**, 491-502.

Ogilvie, M. & the Rare Breeding Birds Panel (1999) Rare breeding birds in the UK in 1997. *British Birds*, **92**, 389-428.

Peach, W.J., Buckland, S.T & Baillie, S.R. (1996) The use of constant effort mist-netting to measure between-year changes in the abundance and productivity of common passerines. *Bird Study*, **43**, 142-156.

Peach, W.J., Baillie, S.R. & Balmer, D.E. (1998) Long-term changes in the abundance of passerines in Britain and Ireland as measured by constant effort mist-netting. *Bird Study*, **45**, 257-275.

Read, J.L., Reid, N. & Venables, W.N. (2000) Which birds are useful bioindicators of mining and grazing impacts in arid South Australia? *Environmental Management*, **26**, 215-232.

Robinson, R.A. & Noble, D.G. (2000) *Regional indicators of wild bird populations: an exploratory study*. BTO Research Report No. 244. British Trust for Ornithology, Thetford.

Sauer J.R. & Geissler, P.H. (1990) Annual indices from route-regression analyses. In: J.R. Sauer & S. Droege, eds., *Survey Designs and Statistical Methods for the Estimation of Avian Population Trends*. US Fish and Wildlife Service Biological Report **90 (1)**.

SAS Institute, Inc. (1996) *SAS/STAT Software: Changes and Enhancements through Release 6.11*. SAS Institute, Inc., Cary, North Carolina.

Sharrock, J.T.R. (1976). *The Atlas of Breeding Birds in Britain and Ireland*. Poyser, London.

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

Sokal, R.R. & Rohlf, F.J. (1995) *Biometry, 2nd Edition.* Freeman, New York, USA.

Steiner, K., Herweg, K. & Dumanski, J. (2000) Practical and cost-effective indicators and procedures for monitoring the impacts of rural development projects on land quality and sustainable land management. *Agriculture, Ecosystems and Environment*, **81**, 147-154.

ter Braak, C.J.F., van Strien, A.J., Meijer, R. & Verstrael, T.J. (1994) Analysis of monitoring data with many missing values: which method? *Bird Numbers 1992: Distribution, Monitoring and Ecological Aspects; Proceedings 12th International Conference of IBCC & EOAC* (eds W. Hagemeijer & T. Verstrael), pp 663-673. SOVON, Beek-Ubbergen, Netherlands.

Thomas, L. (1996) Monitoring long-term population change: why are there so many analysis methods? *Ecology*, **77**, 49-58.

Underhill, L.G. & Prŷs-Jones, R.P. (1994) Index numbers for waterbird populations. I. Review and methodology. *J. Appl. Ecol.*, **31**, 463-480.