BTO Research Report No. 684

Providing Data for Rapid Condition Assessment of Non-Breeding Waterbird SPAs in England: Phase II

Authors

Graham E. Austin, Teresa M. Frost and Kathryn Ross

Report of work carried out by The British Trust for Ornithology

January 2017

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

British Trust for Ornithology

Providing Data for Rapid Condition Assessment of Non-Breeding Waterbird SPAs in England: Phase II

BTO Research Report No 684

Graham E. Austin, Teresa M. Frost and Kathryn Ross

Published in January 2017 by the
British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU, U.K.

ISBN 978-1-908581-73-0

CONTENTS

Page No.

LIST OF TABLES

Page No.

EXECUTIVE SUMMARY

This report explores means by which a standardised trend analysis of data from the Wetland Bird Survey (WeBS) can aid rapid assessment of condition for non-breeding waterbird SPAs in England.

This follows on from pilot work to investigate means by which rapid condition assessment of SPAs in England could be developed, adapting the existing Wetland Bird Survey (WeBS) Alerts approach to fit more closely with Common Standards Monitoring requirements (Austin 2015).

The pilot study identified a suite of potential metrics based on various comparisons of maximum, minimum and mean values between the five-year period upon which notification was based and a five-year assessment period. These comparisons were made when using either raw winter peaks or smooth trends through those peaks. An arbitrary 25% decline in the value of the metric between the notification period and the assessment period was chosen to indicate a decline in apparent condition, 25% being used for consistency with WeBS Alerts.

In this report we have explored the sensitivity of those metrics to determine which provide the most consistent assessments from year to year the aim being to reduce the likelihood of 'false alarms'. This was addressed by creating a time-series of assessments for each metric as if they have been made on an annual basis, the aim being to identify those metrics least likely to give false concern that apparent condition has fallen below a chosen threshold or give false confidence that apparent condition has improved.

A recommendation from the pilot report was to also consider using 'bird-months' - cumulative monthly counts over the winter period – as an alternative to annual peaks. Consequently we have executed all comparisons both when using winter peaks and when using bird months as the measure underpinning the metrics.

These analyses were undertaken for both individual SPA features and also the waterbird assemblage.

A general pattern emerged that although majority of the 'Alerts' metrics were consistent between themselves, marginally more so when based on bird-months rather than peaks, the KNF approach currently in use frequently failed to identify consistent declines in features or assemblage abundance detected by the majority of 'Alerts' metrics .

Our recommendation is that Natural England should adopt an 'Alerts' based metric for rapid condition assessment. If there is a desire to continue using a metric based on peak winter counts then we would recommend adopting that of change in the five-year average of the GAM-smoothed trend through peak winter counts.

However, a switch to using the bird-month equivalent may bring with it the added benefit of identifying a decline in over-winter carrying capacity before this begins to affect assessments based on peak numbers. In this case we would recommend rapid condition assessment be based on the change in the five-year average of the GAM-smoothed trend through cumulative monthly counts over the winter period.

A second major development was to develop a means of identifying any decline in diversity of the waterbird assemblage given that the current approach of comparing number of species recorded during the two five-year periods is a crude comparison most likely to be driven by the incidental occurrence of rarer species.

We therefore developed the Special Protection Area Weighted Indicator Tool (SPAWIT) to allow Natural England staff to visualise trends at the site since the designation period. The tool allows exploration, at the site level, of an indicator formed from the geometric mean of site species indices by the user setting weights by conservation importance at the site and viewing which species are contributing to the site indicator under that weighting scheme.

Three Excel Spreadsheets accompany this report:

1) Results of the analyses of apparent condition based upon the various 'Peak Count' metrics.
    a. Assessment 2016: The assessments using the recommended metrics for the most recent annual assessment (data up to and including those for the winter 2014/15 - the most topical available from WeBS).
    b. ByAssemblage: Rolling assessments through time for assemblage by SPA
    c. ByFeature: Rolling assessments through time for individual features by SPA
    d. Feature reliability
2) Results of the analyses of apparent condition based upon the various 'Bird-month' metrics.
    a. Assessment 2016: The assessments using the recommended metrics for the most recent annual assessment (data up to and including those for the winter 2014/15 - the most topical available from WeBS).
    b. ByAssemblage: Rolling assessments through time for assemblage by SPA
    c. ByFeature: Rolling assessments through time for individual features by SPA
    d. Feature reliability
3) The Special Protection Area Weighted Indicator Tool (SPAWIT).

1. BACKGROUND

In 2015 the BTO published a report to Natural England that constituted pilot work to investigate means by which rapid condition assessment of SPAs in England could be developed, adapting the existing Wetland Bird Survey (WeBS) Alerts approach to fit more closely with Common Standards Monitoring requirements (Austin 2015, BTO Research Report 670). Natural England have requested this follow-up project which is designed to explore outstanding areas and to resolve issues arising from the pilot, making use of the entire WeBS database and existing computer programs developed by BTO on behalf of the WeBS partners. The project is focused on SPAs (and, secondarily, SSSIs) in England.

The primary aim of the project is to develop analysis programmes and tools to interrogate the WeBS database, in order to allow periodic (e.g. every three years) production of data outputs describing changes in bird numbers at SPAs (and potentially SSSIs) in England, for which WeBS data are available.

Specific objectives were as follows:

- Using the WeBS database, test 'sensitivity' of WeBS change metrics developed in the pilot project in order to make firm recommendations to Natural England on the most robust methods to assess changes in bird numbers to deliver information for site condition assessment (section 2);
- Consult with Natural England staff to resolve as many outstanding queries from the pilot project, e.g. issues with citation periods and WeBS data availability, to make comparisons as accurate as possible (section 3);
- Review, develop and test indicators of change for waterbird assemblages using available WeBS data, in order to make firm recommendations to Natural England on the most robust methods to assess changes in assemblages to deliver information for site condition assessment (section 4);
- Produce up to date assessments of change for the most recently available WeBS data period, factoring in the above objectives (separate data output file); and
- Explore the feasibility of assessing changes based on WeBS data summarising 'bird months' rather than peak counts, and the resulting effects on assessment of change, in order to provide information for Natural England to consider further (section 2.2.2).

2. SENSITIVITY OF POTENTIAL METRICS PROPOSED BY THE PILOT STUDY

2.1 Sensitivity of Potential Metrics: Methods

The current method of assessing change, 'Known Natural Fluctuation' (KNF) takes a conservative approach to flagging noteworthy declines by comparing the minimum value during the classification period with the maximum value during the assessment period. As such it will be vulnerable to overlooking serious long-term declines in numbers when faced with an uncharacteristically low value during the classification period or high value during the assessment period. At the species level, KNF can be expected to be least informative for species for which substantial between-year fluctuations in the UK and / or England are typical, such as Bewick's Swan or Bar-tailed Godwit. Two of the metrics proposed by Austin (2015) go part way to addressing this by comparing like with like, i.e. either minimum against minimum or maximum against maximum, and so are less vulnerable to uncharacteristic values pulling in opposite directions. Nonetheless, these metrics are still based on a single value from each period and so are still vulnerable to a single extreme value masking a long-term change. A further metric proposed by Austin (2015) takes this one step further by comparing mean values for each period, an approach which is less likely to overlook a long-term decline as the effect of a single uncharacteristic or extreme value will have less of a disproportionate effect on the resulting assessment. However, these metrics are still sensitive to atypical numbers in just one of year of the five-years assessment period.

Four further metrics were also proposed for consideration by Austin (2015) all of which were based on the modelled trajectory of waterbird numbers rather than the raw data. They are therefore comparing values from the underlying trend modelled using a smoothing spline derived from a Generalized Additive Model (GAM; Hastie & Tibshirani 1990). The degree of smoothing achieved by a GAM is a function of the degrees of freedom made available to the model (if df=1 the GAM fits a linear model; if df=number of years -1 the GAM fits an unsmoothed model). Here we use df=n/3 where n= number of years in the time series. This degree of smoothing is considered to dampen the year to year fluctuation whilst retaining meaningful changes in trajectory of the underlying trend and has been adopted as the standard degree of smoothing by WeBS for representing trends underlying national indices (e.g. Holt *et al*. 2015), for assessing changes in numbers for protected sites for WeBS Alerts (Cook *et al*. 2013), and as used when deriving the Wintering Waterbird Indicators supplied to Defra and the Scottish Government.

Objective 1 required further analysis to that undertaken during the pilot study to test 'sensitivity/stability' of the potential change metrics developed in the pilot. This has been addressed by running analyses that compare the value for each metric derived classification period with the comparative metric for subsequent five-year assessment periods (Table 2.1.i). This was repeated with the five-year assessment period being incremented by one year each time, the first assessment period being that immediately following the five-year period upon which classification was based. Thus the repetitions represent the assessment that would have been made each year based upon those data that would have been available at the time were made up to and including the most recent five-year period for which WeBS data allow comparisons. This was done for both the total waterbird assemblage and individual features of each SPA. Each simulation used all data for the SPA in question up to and including the year immediately following the end of each simulated assessment five-year period – thus ensuring the smoothing would not be influenced by data that would not have been available at the time for any given assessment period. WeBS Alerts reports Alerts Status as assessed for the penultimate winter of the time series and we follow the same practise for the current report. The reasoning behind this is that the end point of data smoothed by a modelling technique such as GAMs is less reliable than for the general time series, as end points are only influenced by data for preceding years whereas mid-points are influenced by prior and post data. Consequently, the smoothed value for a given year does not become stable until subsequent data become available. It is therefore common practise to treat the end point with caution and use the preceding year as that against which to assess change (for example see Table 2.1.ii). WeBS Alerts treats a decline of 25% or greater as cause for triggering a WeBS Alert. In keeping with this, we use a decline of 25% or greater as giving cause for concern for the current report. However, 25% is an arbitrary value that was adopted following a stakeholders workshop during development of WeBS Alerts and a different value could be chosen.

Table 2.1.i Metrics proposed by Austin (2015), based on annual peak counts, that have been assessed for 'sensitivity/stability' (based on an arbitrary 25% decline giving cause for concern).

Metric	Value from classification period	Value from assessment period	Change flagged as noteworthy (giving "cause for concern")
Known Natural Fluctuation (Pk-KNF)	minimum winter peak	maximum Winter Peak	assessment < classification
Δ mean-Pk	mean winter peak	mean winter peak	assessment < 0.75(classification)
Δ min- Pk	minimum winter peak	minimum winter peak	assessment < 0.75(classification)
Δ max-Pk	maximum winter peak	maximum winter peak	assessment < 0.75(classification)
Δ KNF-Pk-GAMs	minimum trend value	maximum trend value	assessment < classification
Δ mean-Pk-GAMs	mean trend value	mean trend value	assessment < 0.75(classification)
Δ min-Pk-GAMs	minimum trend value	minimum trend value	assessment < 0.75(classification)
Δ max-Pk-GAMs	maximum trend value	maximum trend value	assessment < 0.75(classification)
Δ centre-Pk-GAMs	trend value at mid-point	trend value at mid-point	assessment < 0.75(classification)

Table 2.1.ii Example of simulations for a given SPA (Ribble and Alt Estuaries SPA).

Classification Period	Assessment Period (penultimate 5-winter Period of time series)	Earliest Data Available	Latest Data Used
1995/96 - 1999/00	2000/01 - 2004/05	1994/95 - 1998/99	2001/02 - 2005/06
1995/96 - 1999/00	2001/02 - 2005/06	1994/95 - 1998/99	2002/03 - 2006/07
1995/96 - 1999/00	2002/03 - 2006/07	1994/95 - 1998/99	2003/04 - 2007/08
1995/96 - 1999/00	2003/04 - 2007/08	1994/95 - 1998/99	2004/05 - 2008/09
1995/96 - 1999/00	2004/05 - 2008/09	1994/95 - 1998/99	2005/06 - 2009/10
1995/96 - 1999/00	2005/06 - 2009/10	1994/95 - 1998/99	2006/07 - 2010/11
1995/96 - 1999/00	2006/07 - 2010/11	1994/95 - 1998/99	2007/08 - 2011/12
1995/96 - 1999/00	2007/08 - 2011/12	1994/95 - 1998/99	2008/09 - 2012/13
1995/96 - 1999/00	2008/09 - 2012/13	1994/95 - 1998/99	2009/10 - 2013/14

Further to the metrics proposed by Austin (2015), NE requested we followed up the recommendation of considering 'bird-months' (i.e. the cumulative monthly count for a given winter) as an alternative to using the annual peak counts to derive the various metrics. Bird-months as a measurement has the potential to pick up changes in carrying capacity that could be overlooked when using peak counts. For example if resources are depleted more rapidly in some winters compared to others, waterbird abundance over the winter may be expected to decline more in some winters than others even although peak numbers remain unchanged. Metrics based on bird-months therefore have the potential to respond to increasingly unfavourable condition earlier in the annual time series than those based on peak count, in turn allowing more timely investigation into potential causes. Accordingly, all the metrics were also derived using the bird-month equivalent (Table 2.1.iii).

Table 2.1.iii Metrics, equivalent to those listed in Table 2.1.i except based on bird-months rather than annual peaks, that have been assessed for 'sensitivity/stability' (based on an arbitrary 25% decline giving cause for concern).

Metric	Value from classification period	Value from assessment period	Change flagged as noteworthy (giving "cause for concern")
Known Natural Fluctuation (BM-KNF)	minimum winter bird-months	maximum Winter bird-months	assessment < classification
Δ mean-BM	mean winter bird-months	mean winter bird-months	assessment < 0.75(classification)
Δ min-BM	minimum winter bird-months	minimum winter bird-months	assessment < 0.75(classification)
Δ max-BM	maximum winter bird-months	maximum winter bird-months	assessment < 0.75(classification)
Δ BM-KNF-GAM	bird-months trend minimum value	bird-months trend maximum value	assessment < classification
Δ mean-BM–GAM	bird-months trend mean value	bird-months trend mean value	assessment < 0.75(classification)
Δ min-BM-GAM	bird-months trend minimum value	bird-months trend minimum value	assessment < 0.75(classification)
Δ max-BM-GAM	bird-months trend maximum value	bird-months trend maximum value	assessment < 0.75(classification)
Δ centre-BM-GAM	bird-months trend value at mid-point	bird-months trend value at mid-point	assessment < 0.75(classification)

It should be noted when considering WeBS data that not all species of waterbird are monitored with equal reliability at the site level. Coastal waders are generally reliably monitored by WeBS as are the majority of ducks. Aside from Brent Goose which occupy coastal habitats, many notifications for geese relate to numbers roosting at the SPA in question and the majority of individuals will be foraging over the wider countryside during the middle of the day when standard WeBS counts are undertaken. Another group of species less reliably monitored are species that use off-shore habitats such as scoters and Eider; depending on weather conditions and visibility, these may be beyond the range of on-shore observers. Cryptic species such as Bittern and Snipe are also poorly monitored. Although there is no absolute measure of how well a given species is monitored by WeBS some general guidelines can be given (Table 2.2.iv).

Table 2.1.iv Guide to reliability of WeBS monitoring of features of the SPAs included in this report.

Species	WeBS monitoring	Additional Notes
Bewick's Swan	Well monitored	Roost counts used for by WeBS for Ouse and Nene Washes
Whooper Swan	Well monitored	Roost counts used for by WeBS for Ouse Washes

Species	WeBS monitoring	Additional Notes
Bean Goose (Taiga)	Poorly monitored	Generally not on site during WeBS day-time counts
Pink-footed Goose	Poorly monitored	Generally not on site during WeBS day-time counts
White-fronted Goose (European)	Poorly monitored	Generally not on site during WeBS day-time counts
Greylag Goose (Icelandic)	Poorly monitored	Generally not on site during WeBS day-time counts
Barnacle Goose (Svalbard)	Poorly monitored	Generally not on site during WeBS day-time counts
Brent Goose (Dark-bellied)	Well monitored	
Brent Goose (Light-bellied)	Well monitored	
Shelduck	Well monitored	
Wigeon	Well monitored	
Gadwall	Well monitored	
Teal	Well monitored	
Mallard	Well monitored	
Pintail	Well monitored	
Shoveler	Well monitored	
Pochard	Well monitored	
Tufted Duck	Well monitored	
Scaup	Poorly monitored	Substantial proportion occur offshore with many beyond range of shore based counts
Eider	Poorly monitored	Substantial proportion occur offshore with many beyond range of shore based counts
Common Scoter	Poorly monitored	Substantial proportion occur offshore with many beyond range of shore based counts
Velvet Scoter	Poorly monitored	Substantial proportion occur offshore with many beyond range of shore based counts
Goldeneye	Well monitored	
Red-breasted Merganser	Well monitored	
Cormorant	Well monitored	
Bittern	Poorly monitored	Cryptic species, also recorded in numbers to low to support trend analysis
Little Egret	Well monitored	
Little Grebe	Poorly monitored	Cryptic species, also recorded in numbers to low to support trend analysis
Great Crested Grebe	Well monitored	
Slavonian Grebe	Poorly monitored	Substantial proportion occur offshore with many beyond range of shore based counts
Coot	Well monitored	
Oystercatcher	Well monitored	
Avocet	Well monitored	
Ringed Plover	Well monitored	
Golden Plover	Well monitored	
Grey Plover	Well monitored	
Lapwing	Well monitored	
Knot	Well monitored	
Sanderling	Well monitored	
Purple Sandpiper	Well monitored	
Dunlin	Well monitored	
Ruff	Well monitored	
Snipe	Poorly monitored	Cryptic species, also recorded in numbers too low to support trend analysis
Bar-tailed Godwit	Well monitored	
Black-tailed Godwit	Well monitored	
Curlew	Well monitored	
Redshank	Well monitored	
Turnstone	Well monitored	

2.2 Sensitivity of Potential Metrics: Results

The detailed results from these analyses are presented in two Excel Workbooks that accompany this report, one for annual peak derived metrics ('PkAssessmentsThroughTime2014.xlsb') and one for the bird-month derived metrics ('BMAssessmentsThroughTime2014.xlsb'). In each case, results for assemblages are presented in the sheet 'ByAssemblage' and those for individual features for each site in the sheet 'ByFeatures'. The sheets 'StatusChange for Assemblage' and 'StatusChange for ByFeature' provides a cross tabulation (pivot table) that allows the frequency of change in status between consecutive assessments for the assemblage or a given species on a given SPA to be explored. Status is defined here as whether or not a given metric has flagged a decline as noteworthy or not using the arbitrary 25% decline threshold where appropriate (metrics other than KNF).

2.2.1 Annual Peak-based Metrics

A by SPA précis of the results for consecutive assessments of the assemblage (Table 2.2.1.i) indicates that in general there is a high level of agreement between assessments made using the various annual peak based metrics. Of the 55 SPAs that could be assessed the majority were consistently assessed as either giving no cause for concern (36) or giving cause for concern (11) across all metrics. For the remaining eight there is some degree of inconsistency between the metrics although six of those could on balance be classed as giving cause for concern (3) or no cause for concern (3) with only two with a roughly even split such that it is not clear which way to call the result.

Although the pattern is weak, where there are discrepancies in the assessments obtained from the various metrics it is apparent that the KNF metrics often differ from the 'Alerts' (25% decline) metrics. This is not surprising given that the former compare a five-year minimum to a five-year maximum and flag any decline, whereas the latter compare like with like flagging when a threshold percentage decline has been exceeded. However, it is of concern that the KNF metrics fail to respond to consistent year on year declines.

Although it would be possible to work through the same process for the assessments of individual SPA features the resulting tabulations would be somewhat unwieldy and there is no obvious means by which to summarise those results. However, the general impression gained from the Excel spreadsheets is again one of general consistency across all 'Alerts' metrics but with a high frequency of KNF metrics failing to respond to consistent year on year declines.

Table 2.2.1.i: By SPA précis of conclusions drawn for SPA assemblages from metrics based on peak winter counts.

SPA	Assessments for Assemblage	Apparent Condition (of assemblage)
Upper Solway Flats and Marshes	Aside from min-PK, all metrics indicate cause for concern by 2007/08-2011/12.	Poor
Duddon Estuary	All metrics indicate cause for concern for most recent assessment, however prior to that picture is less clear. The GAM-based metrics are in broad agreement in indicating cause for concern since the assessment of 1998/99-2002/03 or thereabouts and intermittently thereafter.	Poor
Morecambe Bay	All metrics are in agreement indicating no cause for concern.	Good

SPA	Assessments for Assemblage	Apparent Condition (of assemblage)
Ribble and Alt Estuaries	Five metrics, KNF-Pk, Mean-Pk, Min-Pk & min-Pk-GAM indicate cause for concern for the most recent one or two assessments whilst KNF-Pk-GAM indicates cause for concern throughout the time-series. Additionally, KNF-Pk also indicated cause for concern up until the assessment for 2003/04-2007/08, after which there was a respite until the most recent assessment.	Inconsistent
Martin Mere	All metrics are in agreement indicating no cause for concern.	Good
Mersey Estuary	All metrics are consistent in indicating cause for concern from the assessment of 2003/04-2007/08 or shortly thereafter. However, the KNF-PK stands alone in indicating no cause for concern since the 2007/08-2011/12 assessment.	Poor
Lindisfarne	All metrics are in agreement indicating no cause for concern.	Good
Holburn Lake and Moss	All metrics are in agreement indicating no cause for concern.	Good
Teesmouth and Cleveland Coast	Max-Pk stand alone in indicating cause for concern in earlier years. Together with KNF-Pk_GAM, Max-Pk-GAM &Cent-Pk-GAM it also indicates cause for concern in the most recent assessments. The remaining metrics are in agreement indicating no cause for concern throughout the time-series.	Inconsistent
Lower Derwent Valley	All metrics are in agreement indicating no cause for concern with the exception of Max-Pk-GAM which indicated cause for concern for two consecutive assessments beginning with the 1999/00-2003/04 assessment.	Good
Humber Estuary	All metrics are in agreement indicating no cause for concern throughout the time-series with the exception of Max-Pk and Cent-Pk-GAM which indicate cause for concern in the most recent one or two assessments respectively.	Good
Northumbria Coast	Max-Pk is exceptional in indicating cause for concern throughout the time series. In general other metrics indicate no cause for concern although Mean-Pk, Max-Pk, mean-Pk-GAM and Cent-Pk-Gam all exhibit between one and three, generally consecutive, assessments indicating cause for concern.	Good
Hornsea Mere	All metrics are in agreement indicating no cause for concern.	Good
Walmore Common	All metrics are in agreement indicating no cause for concern with the exception of the Min-Pk-GAM & Cent-PK-GAM which indicated cause for concern for one or two consecutive assessments near the beginning of the time-series.	Good
The Wash	All metrics are in agreement indicating no cause for concern.	Good
Gibraltar Point	All metrics are in agreement indicating no cause for concern.	Good
Nene Washes	All metrics are in agreement indicating no cause for concern.	Good
Ouse Washes	All metrics are in agreement indicating no cause for concern.	Good
Rutland Water	All metrics are in agreement indicating no cause for concern.	Good
North Norfolk Coast	All metrics are in agreement indicating no cause for concern.	Good
Minsmere-Walberswick	Although in general all metrics indicate no cause for concern, Min-Pk stands out as indicating cause for concern for the first five assessments of the time series and min-Pk-GAM & Cent-Pk-GAM indicate cause for concern for isolated assessments near the beginning of the time-series.	Good
Alde-Ore Estuary	All metrics are in general agreement indicating cause for concern for recent assessments although precise patterns differ. KNF-Pk, Max-Pk & KNF-Pk-GAM stand out as indicating cause for concern throughout the time-series.	Poor

SPA	Assessments for Assemblage	Apparent Condition (of assemblage)
Stour and Orwell	All metrics are in close agreement indicating cause for concern since the 2003/04-2007/08 assessment (or thereabouts). KNF-PK-GAM was the first to indicate cause for concern.	Poor
Hamford Water	Results are somewhat polarized between Mean-Pk, Max-Pk,, KNF-Pk-GAM, Mean_pk_GAM, Max-Pk-_GAM & Cent-Pk-GAM which indicate cause for concern through the time series barring the most recent assessment and the remaining metrics which indicate no cause for concern.	Good for latest assessment (previously Inconsistent)
Abberton Reservoir	Results are inconsistent although the majority of the metrics indicate cause for concern for the 2000/01-2004/05 assessment with some doing so for several assessments prior to that. KNF-Pk and Mean-Pk-GAM however indicate no cause for concern throughout the time-series and Min-PK alone indicates cause for concern for five consecutive assessments starting with that for 2002/03-2006/07.	Good for recent assessments but has varied over time (Inconsistent)
Benfleet and Southend Marshes	All metrics are in agreement indicating no cause for concern.	Good
Breydon Water	All metrics are in agreement indicating no cause for concern.	Good
Dengie (Mid-Essex Coast Phase 1)	All metrics are in agreement indicating no cause for concern.	Good
Colne Estuary (Mid-Esses Coast Phase 2)	All metrics are in broad agreement indicating cause for concern since the 2001/02-2005/06 assessment or from one to three year prior to that however. KNF-Pk, Max-Pk and Max-Pk-GAM indicate no cause for concern for the most recent two or three assessments.	Poor
Crouch and Roach Estuaries (Mid-Essex Coast Phase 3)	All metrics are in agreement indicating no cause for concern.	Good
Blackwater Estuary (Mid-Essex Coast Phase 4)	All metrics are in agreement indicating no cause for concern.	Good
Foulness (Mid-Essex Coast Phase 5)	All metrics are in agreement indicating no cause for concern.	Good
Broadland	All metrics are in agreement indicating no cause for concern.	Good
Deben Estuary	All metrics are in agreement indicating no cause for concern.	Good
Somerset Levels and Moors	All metrics are in agreement indicating no cause for concern.	Good
Chew Valley Lake	All metrics are in agreement indicating no cause for concern throughout the time-series with the exception of KNF-Pk in the first year of the time-series.	Good
Exe Estuary	All metrics are in agreement indicating no cause for concern throughout the time series with the exception KNF-Pk which indicated cause for concern for the five consecutive assessments beginning with that of 2004/05-2008/09 and KNF-Pk-GAM which indicated cause for concern for all assessments since that of 1996/97-2000/01 (barring 1997/98-2001/02) for the most recent seven and nine assessments respectively.	Good
Chesil Beach and The Fleet	Results for the various metrics are somewhat inconsistent. Min-Pk, KNF-Pk-GAM & Min-Pk-GAM alone indicating cause for concern for the most recent two or three assessments although KNF-Pk-GAM has intermittently indicated cause for concern throughout the time-series.	Poor for recent assessments (Inconsistent)

SPA	Assessments for Assemblage	Apparent Condition (of assemblage)
Poole Harbour	All metrics are in broad agreement indicating cause for concern for assessments from the second half of the time-series following a period during which none indicated cause for concern. However, the assessment from which the indication tipped from no cause for concern to cause for concern is inconsistent. KNF-Pk and KNF-Pk-GAM have indicated cause for concern throughout the time series barring the first assessment.	Poor
Tamar Estuaries Complex	Results are highly polarized with KNF-Pk, Min-Pk, KNF-Pk-GAM and Min-Pk-GAM indicating no cause for concern throughout the time-series and the remaining metrics indicating cause for concern consistently from early in the time-series.	Poor (Inconsistent)
Chichester and Langstone Harbours	KNF-Pk-GAM stands out as indicating cause-for-concern throughout the time series whereas, with the exception of Max-Pk-GAM which indicates cause for concern for the last two assessments, none of the other metrics indicate cause for concern for any assessment.	Good
Portsmouth Harbour	There is a broad agreement between metrics other than KNF-Pk in indicating cause for concern for the 1998/99-2002/03 assessment, some also indicating the same one or two years either side. Max-Pk stands alone in indicating cause for concern for the two most recent assessments	Poor
Solent and Southampton Water	Results are somewhat polarized with KNF-Pk & KNF-Pk-GAM indicating cause for concern consistently since the 1999/00-2002/03 assessment. Min-Pk, Cent-Pk and Min-Pk-GAM all exhibit one of more assessments from 2004/05-2008/09 indicating cause for concern but with inconsistency in the number of consecutive assessments before returning to no cause for concern.	Good at present (inconsistent)
Avon Valley	All metrics are in agreement indicating no cause for concern.	Good
The Swale	All metrics are in agreement indicating no cause for concern with the exception of Min-Pk which indicates cause for concern for the most recent assessment.	Good
Thames Estuary and Marshes	All metrics are in agreement indicating no cause for concern.	Good
Medway Estuary and Marshes	All metrics are in agreement indicating cause for concern throughout the time-series.	Poor
Pagham Harbour	All metrics are in agreement indicating no cause for concern.	Good
Thanet Coast and Sandwich Bay	All metrics are in agreement indicating no cause for concern; the rather trivial exceptions being Min-Pk & Cent-Pk for the first assessment of the time-series only.	Good
Dungeness to Pett Level	All metrics are in agreement indicating no cause for concern.	Good
Lee Valley	All metrics are in agreement indicating no cause for concern.	Good
South West London Waterbodies	The results are somewhat polarized with KNF-Pk & Min-Pk indicating no cause for concern whilst the remainder of the metrics are in broad agreement indicating cause for concern consistently through all or most of the time-series although there is some inconsistency as to the first assessment indicating cause for concern.	Poor (Inconsistent)
Dee Estuary	All metrics are in agreement indicating no cause for concern throughout the time series with the exception of KNF-Pk-GAM for the 2006/07-2010/11 assessment only.	Good

SPA	Assessments for Assemblage	Apparent Condition (of assemblage)
Severn Estuary	All metrics are in agreement indicating cause for concern throughout the time-series with the exception of KNF-Pk and KNF-Pk-GAM which indicate cause for concern for the second and first and second assessments only.	Poor
Arun Valley	All metrics are in broad agreement indicating cause for concern from mid-way through the time-series although the precise assessment from which this state persists varies to within a couple of years. However, KNF-Pk and Min-Pk are exceptions in indicating no cause for concern throughout the time series.	Poor

2.2.2 Bird-month based Metrics

A by SPA précis of the results for consecutive assessments for the assemblage (Table 2.2.1.i) indicates that in general there is a high level of agreement between assessments made using the various bird-month based metrics. The results are broadly similar with those obtained when using the annual peak based metrics. Of the 55 SPA that could be assessed the majority were consistently assessed as either giving no cause for concern (37) or giving cause for concern (14) across most metrics. For the remaining four there is some degree of inconsistency between the metrics although two of those could on balance be classed as giving cause for concern, the remaining two having a roughly even split such that it is not clear which way to call the result.

Using the bird-months based metrics has therefore removed much of the inconsistency from half of those SPAs where this was a problem. What it has also done in some cases is polarize the difference between the KNF metrics (comparing minimum during the notification period to maximum during the assessment period) with the other alerts metrics.

As for the annual peak based metrics, whilst it would be possible to work through the same process for the assessments of individual SPA features the resulting tabulations would be unwieldy and there is no obvious means by which to summarise those results statistically. Once again, the general impression gained from the Excel spreadsheets is one of general consistency across all 'Alerts' metrics but with a high frequency of KNF metrics failing to respond to consistent year on year declines.

Table 2.2.1.i: By SPA précis of conclusions drawn for SPA assemblages from metrics based on bird-months (cumulative winter count).

SPA	Assessments for Assemblage	ApparentCondition (of assemblage)
Upper Solway Flats and Marshes	Aside from min-BM, all metrics indicate cause for concern by 2007/08-2011/12. Earliest warning came from ΔKNF-BM-GAM.	Poor
Duddon Estuary	All metrics indicate cause for concern by 2005/06-2008/09. KNF, KNF-BM-GAM, Max-B, Max-BM-GAM Cent-BM-GAM had indicated cause for concern between 6 and 8 years earlier.	Poor
Morecambe Bay	Aside from KNF-BM-GAM which flags cause for concern from 2008/09-2013/14; all metrics indicate there is no cause for concern throughout the time series.	Good

SPA	Assessments for Assemblage	ApparentCondition (of assemblage)
Ribble and Alt Estuaries	Aside from the min-BM & Min-BM-GAM all metrics flag cause for concern since 2008/09-2013/14. However the KNF-BM & KNF-BM-GAM both indicate cause for concern throughout the time series.	Poor (Inconsistent)
Martin Mere	All metrics are in agreement indicating no cause for concern.	Good
Mersey Estuary	All metrics are consistent in indicating cause for concern within a year of the 2003/04-2007/08 assessment period. However, the KNF-BM stands alone in indicating no cause for concern since the 2007/08-2011/12 assessment.	Poor
Lindisfarne	The various metrics are broadly in agreement in indicating cause for concern throughout the time-series with the exception of a respite between the 1998/99-2002/03 assessment period to the 2002/03-2006/07 assessment period, give or take a year. However Max-BM, Max-BM-GAM & KNF-BM-GAM do not identify this respite period.	Poor
Holburn Lake and Moss	All metrics are in agreement indicating no cause for concern.	Good
Teesmouth and Cleveland Coast	KNF-BM & KNF-BM-GAM alone based metrics stand out as indicating cause for concern in earlier years. There is broad agreement across all metrics in indicating cause for concern in the most recent assessments.	Good
Lower Derwent Valley	All metrics are in agreement indicating no cause for concern.	Good
Humber Estuary	General agreement between the various metrics with all indicating cause for concern in the most recent assessment period (2009/10 – 2013/14) with most doing so for one or two years prior to that.	Poor
Northumbria Coast	KNF-BM-GAM stands alone in indicating cause for concern throughout the time-series. Other than Min-BM & Max-BM-GAM there is a general indication of cause for concern for the two to four most recent assessments.	Poor for most recent assessments
Hornsea Mere	All metrics are in agreement indicating no cause for concern.	Good
Walmore Common	All metrics are in agreement indicating no cause for concern with the exception of the Cent-BM-GAM which indicated cause for concern over two consecutive assessment periods in the early 1990s.	Good
The Wash	All metrics are in agreement indicating no cause for concern.	Good
Gibraltar Point	Inconsistency between metrics with assessments rather polarized; five metrics indicating cause for concern throughout most of the time-series, two metrics indicating no cause for concern and two showing fluctuation between cause for concern/no cause for concern.	Inconsistent
Nene Washes	All metrics are in agreement indicating no cause for concern.	Good
Ouse Washes	KNF-BM-GAM stands alone in indicating cause for concern during the most recent three assessment periods and the first eight assessments with a seven-year respite between. All other metrics are in agreement indicating no cause for concern.	Good

SPA	Assessments for Assemblage	ApparentCondition (of assemblage)
Rutland Water	All metrics are in agreement indicating no cause for concern.	Good
North Norfolk Coast	All metrics are in agreement indicating no cause for concern.	Good
Minsmere-Walberswick	All metrics are in agreement indicating no cause for concern.	Good
Alde-Ore Estuary	All metrics are in agreement indicating no cause for concern.	Good
Stour and Orwell	All metrics are in close agreement indicating cause for concern since the 2003/04-2007/08 assessment (or thereabouts). KNF-BM-GAM was the first to indicate cause for concern.	Poor
Hamford Water	All metrics are in agreement indicating no cause for concern.	Good
Abberton Reservoir	Most metrics indicate cause for concern for one or more consecutive assessments centred on the 1999/00-2003/04 assessment. Exceptions are KNF-BM, Min-BM & Min-BM-GAM.	Good for recent assessments
Benfleet and Southend Marshes	Aside from KNF-BM & KNF-BM-GAM which indicate cause for concern for the most recent two assessments, all metrics are in agreement indicating no cause for concern throughout the time-series.	Good
Breydon Water	Aside from three metrics Mean-BM, Max-BM & Cent-BM-GAM which indicate cause for concern for the most recent assessment and, in the case of the latter, for two assessments prior to that, all metrics are in agreement in indicating no cause for concern throughout the time-series.	Good
Dengie (Mid-Essex Coast Phase 1)	All metrics are in agreement indicating no cause for concern.	Good
Colne Estuary (Mid-Esses Coast Phase 2)	All metrics are in broad agreement indicating cause for concern since the 2002/03-2006/07 assessment or from one to three year prior to that. The earliest indications of cause for concern come from Min-BM & Min-BM-GAM.	Poor
Crouch and Roach Estuaries (Mid-Essex Coast Phase 3)	KNF-BM-GAM stands alone in indicating cause for concern for any period in the time-series, this being for four consecutive assessments beginning with that for 1996/97-2000/01.	Good
Blackwater Estuary (Mid-Essex Coast Phase 4)	All metrics are in agreement indicating no cause for concern.	Good
Foulness (Mid-Essex Coast Phase 5)	All metrics are in agreement indicating no cause for concern.	Good
Broadland	All metrics are in agreement indicating no cause for concern.	Good
Deben Estuary	KNF-BM & KNF-BM-GAM alone indicate cause for concern for any assessments throughout the time-series.	Good
Somerset Levels and Moors	All metrics are in agreement indicating no cause for concern.	Good
Chew Valley Lake	All metrics are in agreement indicating no cause for concern throughout the time-series with the exception of the raw maximum based metric which indicated cause for concern for first assessment of the time-series in 1996/96-2000/01.	Good

SPA	Assessments for Assemblage	ApparentCondition (of assemblage)
Exe Estuary	All metrics are in agreement indicating no cause for concern throughout the time series with the exception KNF-BM & KNF-BM-GAM which indicated cause for concern for the most recent seven and nine assessments respectively.	Good
Chesil Beach and The Fleet	All metrics are in agreement indicating no cause for concern throughout the time series with the exception of KNF-BM & KNF-BM-GAM which indicated cause for concern throughout the time series.	Good
Poole Harbour	All metrics are in agreement indicating cause for concern since the assessment of 2003/04-2007/08 or thereabouts. Exceptions are the KNF-BM & KNF-BM-GAM that have indicated cause for concern throughout the time series cause for concern for the most recent seven to nine assessments.	Poor
Tamar Estuaries Complex	All metrics are in agreement indicating cause for concern since the assessment of 1997-99/04-2001/02 or within one or two assessments thereafter.	Poor
Chichester and Langstone Harbours	Results are somewhat polarized between KNF-BM & KNF-BM-GAM both of which indicate cause for concern since the 2003/04-2007/08 assessment or throughout the time series respectively, and the remaining metrics all of which indicate no cause for concern throughout the time-series (with the exception of two isolated assessments neither of which is sustained) .	Good
Portsmouth Harbour	There is inconsistency between metrics although five indicate cause for concern over one of more assessments centred on the 1997/98-2001/02 assessments and four indicate cause for concern for between one and three most recent assessments.	Inconsistent
Solent and Southampton Water	There is inconsistency between metrics although most indicate cause for concern at least once during the most recent four assessments before which they indicate no cause for concern. KNF-BM-GAM stands out as indicating cause for concern throughout the time series.	Poor for most recent assessments (inconsistent)
Avon Valley	All metrics are in agreement indicating no cause for concern.	Good
The Swale	All metrics are in agreement indicating no cause for concern.	Good
Thames Estuary and Marshes	All metrics are in agreement indicating no cause for concern.	Good
Medway Estuary and Marshes	All metrics are in broad agreement indicating cause for concern since 1998/99-2002/03 or one or two years prior.	Poor
Pagham Harbour	All metrics are in agreement indicating no cause for concern with the exception of KNF-BM-GAM indicating cause for concern in the most recent assessment.	Good
Thanet Coast and Sandwich Bay	All metrics are in agreement indicating no cause for concern.	Good
Dungeness to Pett Level	All metrics are in agreement indicating no cause for concern.	Good
Lee Valley	All metrics are in agreement indicating no cause for concern.	Good

SPA	Assessments for Assemblage	ApparentCondition (of assemblage)
South West London Waterbodies	All metrics are in broad agreement indicating cause for concern since 2000/01-2004/05 or from the start of the time series three years previous.	Poor
Dee Estuary	All metrics are in broad agreement indicating no cause for concern throughout the time series with the exception of KNF-BM and KNF-BM-GAM which indicated cause for concern for the 2007/08-2010/11 assessment and in the case of the latter for the two prior assessments also.	Good
Severn Estuary	All metrics are in broad agreement indicating cause for concern throughout most of the time-series.	Poor
Arun Valley	All metrics are in broad agreement indicating cause for concern throughout most of the time-series.	Poor

2.3 Sensitivity of Potential Metrics: Discussion

Although there is no absolute right or wrong in the choice of the best metric to characterise the change in numbers of a given species or the SPA assemblage, the preferred metric would balance the need to flag noteworthy declines in numbers since notification i.e. those of a magnitude that would give cause for concern with the property of remaining relatively stable with the addition of a new year's worth of data to the time series. A metric that toggles regularly from one year to the next in response to minor fluctuations in numbers runs the risk of being misleading, especially if assessments were to be made periodically rather than annually, or undermining confidence in its meaningfulness if seen to support a different conclusion with each passing year.

When aiming to get an overview of the relative performance of the various metrics, it can be difficult to see past the detail when considering individual features on each site. One way to get a general impression of relative performance of the metrics is to consider the overall frequency of changes in status of the metrics when taken across all SPAs for all features (see Excel sheet 'StatusChange for ByFeature'). KNF recorded substantially fewer changes in status (156) than any of the other metrics (between 291 and 420) when using the annual peak count based metrics. However, it would appear that this comes about as a result of the undesirable failure of KNF to flag noteworthy declines rather than what would be a desirable property of stability in the face of minor fluctuations. The remaining metrics all recorded higher and broadly similar frequencies of changes in status. It would appear that these higher frequencies come about because of the desirable property of not failing to flag noteworthy declines rather than instability, the general impression being one of consistency for a series of incremental assessments following a change in status. However, looking within this second group of metrics for the one with the lowest frequency, and hence the one least likely to flag assessments as noteworthy only to revert the following year, that based on the comparison of five-mean values based on the GAM modelled trend exhibits greatest stability (291 changes in status). However, when using the bird-months based metrics KNF was more comparable with the other metrics.

Comparing this with the bird-month based metrics is interesting in that using this approach the KNF equivalent performs similarly to the other metrics and with 298 changes in status falls within the range of the other metrics (278 to 366). Again the five-year mean GAM metric mean-BM-GAM appears to be a good option although the KNF-BM-GAM is marginally better.

One minor disadvantage shared by all metrics based on the smoothed trend is the need to back date the assessment five-year period by one year due to the inherent instability of the end point in a smoothed spline (as modelled value is only influenced by prior data whereas the following year it will also be influenced by post data). However, it is considered that this disadvantage is easily outweighed by the insensitivity of these metrics to uncharacteristic and extreme values.

2.4 Sensitivity of Potential Metrics: Conclusions and Recommendations

We would recommend moving to the 'Alerts' concept metrics given that an uncharacteristically low annual peak during the five years period for notification or a single high count in the five-year assessment period can mean condition may be assessed to be favourable despite an ongoing downward trend in numbers. Although there is little to choose between the various 'Alerts' concept metrics, those based on cumulative monthly counts rather than winter peak appear to exhibit greater consistency between consecutive assessments (and each other). Thereafter, the metric based on the five-year mean of GAM smoothed annual values would be the one that intuitively would be the most robust against winters of uncharacteristically low or high waterbird abundance. We therefore conclude that the metric that best achieves the balance between flagging noteworthy declines in numbers without being overly sensitive to minor fluctuations is that based on comparison between the five-year means of the GAM-smoothed cumulative winter count (mean-BM-GAM).

If there is a desire to retain the use of annual peak based metrics rather than bird-month based metrics then the equivalent five-year means of the GAM-smoothed peak winter count (mean-Pk-GAM) would be the one recommended.

We provide assessments for 55 SPAs based on WeBS data up to and including those for the 2014/15 winter, detailing assessment status as of the five-winter period 2009/10 to 2013/14 (Excel Workbook sheets 'Assessments2016.xlsb').

Hereafter we suggest new assessments could be supplied to Natural England on an ongoing three-year schedule although annually revisions could be made available to Natural England on request.

3. OUTSTANDING ISSUES

3.1 Outstanding Issues: Methods

Objective 2 required a list of issues from the pilot project to be compiled and provided to NE. These included the following clarifications:

Confirmation of classification periods for nine sites, documentation for which were unavailable from Natural England at the time of the pilot report. The pilot report had therefore used classification periods as documented by JNCC. This is of concern because classification periods adhered to by Natural England and those documented by JNCC do not correspond for all SPAs. A list of these sites was supplied to Natural England upon commencement of this project and Natural England were tasked with confirming or supplying revised classification periods for those sites.

Six SPAs, all of which have waterbird features and which were listed in the original NE list of SPAs to consider, were not included in the pilot analysis as they have not previously been assessed by WeBS Alerts. However, unlike the pilot study, the current work did not rely on data previously compiled as part of the WeBS Alerts process and did not require direct comparisons of results with those from WeBS Alerts. Consequently, we reconsidered whether WeBS held sufficient data to include such sites in the current analysis.

3.2 Outstanding Issues: Results

Where possible Natural England provided confirmation of classification periods of those SPAs for which the pilot study had followed JNCC classification periods in the absence of periods documented by Natural England (Table 3.2.i). Confirmation by NE for six SPAs remain outstanding at the time of this report.

Table 3.2.i Status of clarification of classifications periods as documented by Natural England for which JNCC documented periods were used for the pilot study.

SPA	Classification period confirmation by Natural England	Classification period as documented by JNCC
Upper Solway Flats and Marshes	Updated to 1986/87 – 1990/91	1991/92 - 1995/96
Martin Mere	Awaiting clarification	1991/92 - 1995/96
Holburn Lake and Moss	Awaiting clarification	1991/92 - 1995/96
The Wash	Awaiting clarification	1992/92 - 1995/96
North Norfolk Coast	Awaiting clarification	1991/92 - 1995/96
Chew Valley Lake	Awaiting clarification	1991/92 - 1995/96
Chesil Beach and The Fleet	Confirmed as 1991/92 – 1995/96	1991/92 - 1995/96
Chichester and Langstone Harbours	Updated to 1992/93 – 1996/97	1991/92 - 1995/96
Pagham Harbour	Awaiting clarification	1991/92 - 1995/96

The six SPAs, all with waterbird features but not included in the pilot analysis have been reassessed for inclusion in future reporting (Table 3.2.ii). Of these, three sites are not monitored by WeBS and so will not be reported in the future. The remaining three sites are currently flagged in the WeBS database as having insufficient data for trend analysis to support WeBS Alerts and so had been automatically rejected by the analysis program during previous development work. For the current report these SPAs were not excluded from the analysis but nonetheless dropped out. Additionally, the Mersey Narrows and North Wirral Foreshore SPA was not assessed because at the time of this report the five-year periods of notification and assessment would overlap. Also there are some technical issues to overcome to separate WeBS data for the Mersey Narrows from the WeBS Mersey Estuary (does not affect The Mersey Estuary SPA) and data from the North Wirral Foreshore from WeBS Dee Estuary data.

Table 3.2.ii SPAs with waterbird features for which analyses were not undertaken for the current analysis, but have been reconsidered for inclusion in future reporting.

SPA	Reason for exclusion from current analysis
Stodmarsh	Flagged as "insufficient data for trend analysis" in WeBS database". Attempts were made to include this SPA in the analysis but the site dropped out of the analysis
Upper Nene Valley Gravel Pits	Flagged as "insufficient data for trend analysis" in WeBS database". Attempts were made to include this SPA in the analysis but the site dropped out of the analysis
Marazion Marsh	Flagged as "insufficient data for trend analysis" in WeBS database". Attempts to include this SPA in the analysis but the site dropped out of the analysis. Note features are Aquatic Warbler (not monitored by WeBS) and Bittern (occurs in numbers too low for trend analysis. Analysis will be for assemblage only
Liverpool Bay	Offshore areas not monitored by WeBS
Outer Thames Estuary	Offshore areas not monitored by WeBS
Coquet Island	Not monitored by WeBS (WeBS database contains data for a single visit)

3.3 Outstanding Issues: Conclusions and Recommendations

With regard to those SPAs for which no Natural England documentation of classification periods were available for the pilot study and for which classification periods were obtained instead from the JNCC web site, one has been confirmed, two revised accordingly and six await clarification from Natural England. Where available the revised classification periods were used for all analyses undertaken for this report.

It is recommended that Natural England work towards confirming or updating the classification periods for the remaining six and advise BTO of these when available. In the interim, future reporting would continue to those classification periods obtained from JNCC. Source of classification will continue to be included in all reporting until such time that all follow Natural England documentation.

With regard to the three SPAs excluded from the current analysis, BTO are investigating further as to why they continue to drop out of the analysis.

4. ASSESSING CHANGES IN SPECIES DIVERSITY

4.1 Assessing Changes in Species Diversity: Background

Objective 3 seeks to explore means of assessing changes in species diversity between the classification period and the assessment period.

The most simplistic method to assess species diversity/richness compares only the total number of waterbird species recorded during the classification period with the total number of species recorded during the assessment period. As such it fails to capture any element of the relative conservation importance of the species concerned, or the fact that within a multi-species assemblage not all species will make an equal abundance contribution. For example, the loss of a species of high conservation importance can be balanced by the arrival of a species of low conservation importance or even the chance recording of a vagrant species. Austin (2015) explored the possibility of refining this approach by ensuring rare and vagrant species were not taken into consideration unless of conservation importance but concluded that the overall concept remained too simplistic to inform conservation assessments in a meaningful way.

Options for a new approach to assessing changes in species diversity were discussed at a meeting between BTO and Natural England (16th Feb 2016). Natural England had requested additional context for rapid assessment of SPAs using species aggregation indices over the time period from classification to present for each SPA, weighted according to conservation concern. Accordingly, Natural England proposed a means of objectively weighting species based on various assessments of conservation status. An action point from that meeting was that BTO would explore how conservation importance-weighted multi-species indices (or indicators) could be developed to inform the rapid assessment process. Initially it was envisaged that the species weightings would be fixed values; however, during development it became apparent that a more flexible approach that would allow Natural England staff to modify weighting to take into consideration local issues and conservation objectives would enhance the indicator.

4.2 Assessing Changes in Species Diversity: Methods

The SPA Weighted Indicator Tool (SPAWIT) is an Excel™ workbook which calculates the weighted geometric mean of species indices for an SPA. From WeBS data for each SPA individual species indices $\mathrm{I_y}$ are supplied:

$\mathrm{I_{sy}} = \frac{N_{sy}}{N_{sb}}$ where $\mathrm{I_{sy}}$is the single-species index for species s in year y,

$\mathrm{N_{sy}}$is the bird-month for species s in year y
i.e. sum of the number of species s counted (or imputed) over winter WeBS counts in year y ,

and $\mathrm{N_{sb}}$is the bird-month for species s in year b
i.e. sum of the number of species s counted (or imputed) over winter WeBS counts in the base year b.

The base year b is the first year of the classification period for the SPA.

Then the weighted geometric mean $\mathrm{G_y}$ is calculated, where

$$\mathrm{G_y} = \exp\left(\frac{\sum_{\mathrm{s}=1}^{\mathrm{n}} \mathrm{w_s} \ln \mathrm{I_{sy}}}{\sum_{\mathrm{s}=1}^{\mathrm{n}} \mathrm{w_s}}\right)$$

where G_y is the multi-species indicator in year y and w_s is the weight for species s.

Where N_{sy} is zero because the species has not yet colonised the site, or there is no available data, the species is not included in the calculation of G_y. In the first year that a new species occurs, its index I_{sy} is set to be equal to G_y (calculated excluding that species) and this becomes the base year for that particular species. If N_{sy} is zero when it has been non-zero in previous years, it is substituted with 0.9 before calculating I_{sy}.

The weights w_s are set by the user by choosing values for scoring species classifications (for certain national or international legislation, and usually where the site has nationally important numbers of the species at classification) and the status of the species of the site (proportion of abundance at classification). The user can set the thresholds for what weights should be used for species according to the scores and also the values of the weights. Table 4.2.i shows the default scores and thresholds used to decide weightings in SPAWIT. Species with scores above the amber threshold are given a weight w_s of 1 by default and species with scores above the red threshold are given a weight w_s of 2 by default. All other species are weighted zero and do not contribute to G_y.

Table 4.2.i Default scores in SPAWIT

Red Threshold	32
Amber Threshold	8

% contribution to SPA's assemblage at classification or best earliest evidence		>25%	10-25%	5-10%	1-5%	<1%
ranked species	**Score**	25	10	5	1	0.5
>1% numbers Qualifying features (Stage 1.1, Annex 1)	32	800	320	160	32	16
>1% numbers Qualifying features (Stage 1.2, 'migratory waterbird' non-Annex 1)	32	800	320	160	32	16
SPA assemblage named component	16	400	160	80	16	8
5 Year Mean Peak (2010/11-2014/15) >2,000	8	200	80	40	8	4
nationally important numbers (5 Year Mean Peak (2010/11 - 2014/15)) >1%	8	200	80	40	8	4
BirdLife International/ IUCN Endangered	8	200	80	40	8	4
BirdLife International/ IUCN Vulnerable	4	100	40	20	4	2
BirdLife International/ IUCN Near Threatened	2	50	20	10	2	1
Section 41 ex-BAP species	2	50	20	10	2	1
Red-listed in BOCC 4	2	50	20	10	2	1
Amber-listed in BOCC4	1	25	10	5	1	0.5

4.3 Assessing Changes in Species Diversity: Discussion and Conclusions

Choosing an aggregated species index or indicator must take into account the expressed purpose of the measure of biodiversity. This could be to measure species diversity (e.g. Simpson and Shannon indices which relate to the relative abundance of species in a community) or functional diversity (measuring the diversity of species trait groups, rather than presence or abundance of individual species). For examples using WeBS data of species richness and functional diversity metrics of wader communities in 100 British estuaries, see Méndez *et al*. (2012). The UK Wintering Waterbird Indicator uses the approach of a geometric mean of species indices to summarise overall trends in species populations (Austin *et al*., 2007).

In this case, the desired outcome is to measure the effectiveness of the SPA in protecting the species and habitat it was put in place for. Options for an indicator for English SPAs were discussed by BTO and Natural England on 16th February 2016 and it was agreed the approach of a geometric mean of individual species indices weighted by conservation value was appropriate in this application. Natural England supplied the list of classifications requiring consideration and BTO collated information for these based on available information from the JNCC website/NBN Gateway, Birds of Conservation Concern 4 and WeBS data on nationally important thresholds for species. Lists of SPA named component species had previously been collated for Austin (2015).

Buckland *et al*. (2011) demonstrated that geometric means of individual indices have several useful properties: the index reflects trends in both abundance and species evenness, and is not prone to bias when detectability varies between species. If all species are declining at the same rate, a geometric mean index will show a decrease whereas a Shannon or Simpson index would not.

Geometric mean indices work on a multiplicative scale, so an increase of 50% in an uncommon species has the same impact as a 50% increase in a common species. Local species extinctions and colonisations such as Little Egret cause issues particularly in calculating the species index in the case where the number of birds in the base year N_{sy}; and in calculating the geometric mean, which can only average over positive values. A mathematically similar problem arises with species that which were not surveyed for the whole time span of the index values even when it can be assumed they were present: for example Cormorant has only been included in WeBS surveys since 1986/7. Sequential additions of species with different base years therefore follow the method recommended by Noble *et al*. (2004) and used in the wintering waterbird indicator (Austin *et al*., 2007). Effectively all index values for a species joining the indicator at a point later than the base year are scaled to the overall indicator value for that year, so that new species do not influence the indicator until after the year in which they are introduced.

Severe declines and very large increases in a single species can have a large influence on a multi-species geometric mean index, particularly when only a small number of species are used. To combat this Noble *et al*. (2004) recommended the following:

1. Each species population index (set to 1 in the base year) should have a lower threshold value such that any index below it is replaced with the lower threshold. A threshold of 0.01 was suggested.
2. Each species population index (set to 1 in the base year) should have an upper threshold value such that any index value above it is replaced with the upper threshold. A threshold of 100 was suggested.
3. Minimum thresholds for population size could be set so that species only enter the multi-species index at a minimum number of individuals.
4. A minimum of 10 species should be included, with a preference for at least 15.

Points (1) and (2) are adopted here, with the modification of raising the upper threshold from 100 to 500 to allow for natural fluctuation in flocking species at the site level. For point (3), the weighting system which takes into account the proportion of the species in the classification assemblage sets the weight to 0 for most rare species at a site. Because the SPA Waterbird Index tool is working at a site level, only a small number of species are used to calculate the index at many sites, and so with respect to point (4) caution is advised in interpreting the index: a sharp decline may be caused by non-site specific factors affecting a single influential species. To assist with interpretation of the SPA multi-species index it is suggested that the results be assessed together with:

- Experimenting with the conservation value weights to include and exclude groups of species (and taking note of the species included, listed with the results in SPAWIT results);
- the other SPA metrics provided by BTO (assemblage plots of total species maxima per year; species richness and abundance metric)
- the latest WeBS Alerts for the SPA (Cook *et al.*, 2013).

References

Austin, G. E. (2015). *The Use of Wetland Bird Survey (WeBS) Data for Rapid Condition Assessment of Non-Breeding Waterbird SPAs in England*. BTO Research Report No. 670. British Trust for Ornithology, Thetford.

Austin, G. E., Banks, A. N., & Rehfisch, M. M. (2007). *The UK Wintering Waterbird Indicator : New developments towards a more comprehensive assessment of trends in the wintering waterbird assemblage*. BTO Research Report No. 479. British Trust for Ornithology, Thetford.

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

Cook, A.S.C.P., Barimore, C., Holt, C.A., Read, W.J. and Austin, G.E. (2013). *Wetland Bird Survey Alerts 2009/2010: Changes in numbers of wintering waterbirds in the Constituent Countries of the United Kingdom, Special Protection Areas (SPAs) and Sites of Special Scientific Interest (SSSIs)*. BTO Research Report 641. British Trust for Ornithology, Thetford. http://www.bto.org/volunteer-surveys/webs/publications/webs-annual-report

Hastie, T.J., Tibshirani, R.J., 1990. *Generalized Additive Models*. Chapman & Hall, London.

Méndez, V., Gill, J. A., Burton, N. H. K., Austin, G. E., Petchey, O. L., & Davies, R. G. (2012). Functional diversity across space and time: Trends in wader communities on British estuaries. *Diversity and Distributions*, *18*(4), 356–365. http://doi.org/10.1111/j.1472-4642.2011.00868.x

Noble, D. G., Newson, S. E., & Gregory, R. D. (2004). *Approaches to dealing with disappearing and invasive species in the UK's indicators of wild bird populations.* BTO and RSPB report to Defra. British Trust for Ornithology, Thetford.