

BTO Research Report 578

Analyses to Assess the Effectiveness of Modifications to the BBS Sampling Design

Authors Freeman, S.N. and Noble, D.G.

A report by the British Trust for Ornithology,

February 2006

© British Trust for Ornithology

British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU Registered Charity No: England and Wales 216652 Scotland SC039193

Authors Freeman, S.N. and Noble, D.G.

Analyses to Assess the Effectiveness of Modifications to the BBS Sampling Design

BTO Research Report No. 578

Published in February 2011 by the British Trust for Ornithology The Nunnery, Thetford, Norfolk, IP24 2PU,UK

ISBN No. 978-1-906204-87-7

© British Trust for Ornithology

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form, or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publishers.

CONTENTS

BTO Research Report No. 578 February 2006

LIST OF TABLES

Page No.

BTO Research Report No. 578 February 2006

LIST OF FIGURES

BTO Research Report No. 578 February 2006

APPENDIX

1. EXECUTIVE SUMMARY

- 1. The Breeding Bird Survey (BBS) began in 1994 and is the annual, large-scale monitoring scheme for common terrestrial breeding birds in the United Kingdom, allowing continuous monitoring of population levels. However, certain regions and habitats remain less wellcovered, such as uplands, and effective monitoring of breeding birds in these areas requires an enhanced form of the BBS survey and/or additional, professional fieldwork.
- 2. Here we use simulations to compare power to detect long-term declines between the current BBS and eight alternative ways that the current design might be extended within upland habitats. Existing BBS data from upland squares across the UK were used to derive the parameters for these analyses. We fitted the standard site and year BBS model for a selection of species found in upland areas to get estimates and variances of abundance, which we applied at the start of our simulations. Approximate power was calculated by simulating large numbers of artificial survey data sets. The model was then fitted to each artificial data set, the significance of the variation in year effects evaluated via a standard likelihood-ratio test, and the number of significant outcomes as a proportion of the total replicates adopted as an approximate measure of power.
- 3. We assumed that the regular BBS will continue in its present form as a baseline against which to compare various forms of augmentation. Our power calculations focus on the ability of any enhanced scheme to detect a rate of decline by 1.1% per year over a period of 25 years (a 25% decline). We considered nine different means by which coverage of upland sites might be surveyed: (a) continuation of the BBS in its current form, (b) fourfold increase in the number of (random) squares covered $(= 4 \times$ current survey effort), (c) twofold increase in the number of (random) squares covered $($ = doubling of current survey effort), (d) number of sites visited remains the same as the current BBS, but each surveyor also covers an additional random, adjacent square (= doubling of current survey effort), (e) number of sites visited remains the same as the current BBS, but 50% of surveyors each cover two additional adjacent squares $($ = doubling of current survey effort), (f) continuation of the BBS in its current form, but with an additional 300% increase in squares every third year ($=$ doubling of current survey effort), (g) continuation of the BBS in its current form, but with an additional 600% increase in squares every sixth year (= doubling of current survey effort), (h) a continuation of the BBS at current levels, plus an additional survey three times greater (in terms of sites) carried out every third year. Each surveyor covers 2 adjacent 1 km. squares (= 3 x current survey effort) and (i) as above, but with each site in the additional survey extended to cover 3 adjacent 1 km Squares ($= 4$ x current survey effort).
- 4. Power analyses for BBS at the GB level reveal that for all but three of the 13 target species the existing survey will detect the designated decline with >95% power. The extent of the improvement in power under augmented schemes b-e is close to 100% power for the Common Sandpiper. For the other species scheme (b) is better than (c) but while this difference is substantial for Ring Ouzel, even (c) provides > 90% power for Dipper. Results for schemes (d) and (e) are broadly comparable in magnitude to (c). Schemes (f) and (g) raise the power to 81% and 84% respectively for the Ring Ouzel; for the other species each results in a figure $> 95\%$. Scheme (h) and (i) produce an almost total detection of a significant decline.
- 5. At the current level of coverage in England, six species (Common Sandpiper, Dipper, Ring Ouzel, Kestrel, Siskin and Tree Pipit) fall short of the 95% level of power. Scheme (b) was the most beneficial; under a four-fold increase in the number of squares visited, only Common Sandpiper and Dipper still fall short of 90% power. For these species and Ring Ouzel the gain under (b) is considerably greater than that under the existing regime, which gives power of only around 40% or less.
- 6. For Scotland, six species achieve 95% power at the current BBS level. Augmented analyses were carried out for the remaining seven (Common Sandpiper, Dipper, Kestrel, Ring Ouzel, Siskin, Snipe and Tree Pipit). Schemes (b-e) all produce power well over 90% for Common Sandpiper, Hooded Crow, Red Grouse, Siskin and Snipe. At the opposite extreme, none of the investigated changes are predicted to raise power above even 25% for Ring Ouzel. As with previous analyses, schemes (c-e) tend to produce very similar results, none having a marked and consistent advantage.
- 7. Only Song Thrush shows >95% power to detect the decline in Wales, thus augmented analyses were conducted for the remaining species. Increasing the scale of the BBS even twofold provided almost 100% power for Pied Wagtail and Tree Pipit. The four-fold increase of scheme (b) is however required to raise power above 50% for species such as Kestrel and Snipe, but the relatively limited coverage in Wales means that even an increase of this scale leaves power below 40% for Common Sandpiper and Ring Ouzel**.**
- 8. The power of the current BBS to detect a 25% decline was less than 95% for Curlew, Dipper, Hooded Crow, Kestrel, Pied Wagtail, Snipe and Song Thrush in Northern Ireland. For Dipper and Kestrel, none of the augmented scenarios resulted in powers greater than 35%. For Hooded Crow and Song Thrush (with powers of 94% and 84% respectively under the current sampling regime), all augmentations yielded very high power. For Curlew, some of scenarios 'c' to 'g' resulted in estimates of power of >80% but for Pied Wagtail and Snipe, the power achieved from those scenarios was between 56% and 78%.
- 9. Scenario (a) achieves >80% power for 64% of the species. Almost a half of species achieve this in England, 39% for Scotland and Wales and 25% for Northern Ireland. Scenario (d), where each surveyor covers an additional random, adjacent square increased the amount of species achieving >80% power in all cases, with Wales, Scotland, England and GB all achieving this for 50% or more of the species. For Scenario (f), with an additional 300% increase in squares every third year, the percentage of species achieving >80 power is just slightly higher to that for scenario (d).
- 10. Although scheme (b), that quadrupled survey effort, was the scheme that resulted in the largest increase in power, other options also provided good increases in power, most notably the two options (f) and (g), involving a 3-fold and 6-fold increase in coverage every third or sixth year respectively. Other than for (a) and (b), there was relatively little difference between the increase in power provided by the different options. However, scenarios involving increasing effort through additional adjacent squares rather than additional random squares would require far less effort from fieldworkers in terms of travel time so are more efficient.

2. INTRODUCTION

The Breeding Bird Survey (BBS) began in 1994, initially concurrent with but eventually superseding the Common Birds Census in 2000 as the sole annual, large-scale monitoring scheme for common terrestrial breeding birds in the United Kingdom (Freeman *et al.* 2007). The BBS has the advantage of a formally stratified and randomised design, and has a wider geographical coverage and a more representative sample of habitats than its predecessor, which was largely concentrated in South-Eastern England where the potential volunteer base is greatest.

The greater range of habitats covered by the BBS raises the opportunity, for the first time, of continuous monitoring of population levels in regions and habitats for which this has not been previously practical. However, certain regions and habitats remain less well-covered than those in locations more easily accessible to greater numbers of non-professional surveyors. Upland regions are a good example: though the area of Great Britain in CEH landclasses (Haines-Young *et al*. 2000) 17- 24, and 28-32 ('True' and 'Marginal' uplands) is vast, 72,891 1 km. squares falling into these categories, yet the numbers of BBS squares covered by the BBS volunteers from these areas is disproportionately small. Effective monitoring of breeding birds in these habitats probably requires an enhanced form of the BBS survey and/or additional, professional fieldwork surveys carried out in a manner compatible for combination with the data resource already available. This is the subject of this report.

3. METHODOLOGY

3.1 Models for BBS data

We used existing BBS data from upland squares across the UK to derive the parameters for the analyses. Within Great Britain, upland squares were defined as those categorised in the CEH landclass as true or marginal upland (23 and 24), but for Northern Ireland it was necessary to devise a system for identifying upland squares using CEH land-cover data. Based on previous comparisons of the distribution of land-class and these land-cover measures for Great Britain where both are recorded (Newson, pers. comm.) Northern Ireland squares comprising greater than 50% of upland land-cover were categorised as upland.

A breakdown of the estimated numbers (mean number surveyed per year, excluding 2001 for which there was very poor coverage due to Foot and Mouth Disease) of upland squares covered in Great Britain and Northern Ireland (the latter excluding 1994 & 1995) is:

We adopt the basic model structure employed since the start of the BBS as a national monitoring scheme. The maximum count at the ith square in year j is assumed to follow a Poisson distribution with mean λ_{ii} , where

(1)
$$
\log (\lambda_{ij}) = S_i + Y_j
$$

a (log-)linear combination of simple site (S) and year (Y) effects. The year effects are employed as annual indices of abundance, since the additive nature of the model implies that the expected counts on each site, though they may differ with a year, all change proportionately between any two years. For parameter identifiability, an additional constraint to the model (1) is required. Here we arbitrarily set the year effect in the final year to zero. This permits ready evaluation of the proportional change in numbers at the end of the series relative to any specified earlier year. A further consequence of the constraint is that the expected counts at each site in the final, most recent year, is given simply by the estimates of the site effects (appropriately transformed). We begin the analyses by fitting model (1) for a selection of species found in upland areas and examining the distribution of these site effects (Table 1) to get estimates of abundance, which we then assume to apply at the start of the simulations discussed in the following section. It is assumed that at the start of the period, squares are drawn at random from an infinite population of such squares, whose 'true' site effects are normally distributed with mean μ and variance σ^2 equal to the sample values of Table 1.

Table 1. Average number of squares (rounded to integer values) in true or marginal upland (in Great Britain) recording each species per year, along with the mean and variance of site effects from the fitted Generalized Linear Model. The site effects are on a natural log scale as specified in equation 1.

Approximate power can then be calculated by simulating large numbers of artificial survey data, assuming equation (1) is the correct underlying form for the data and matched to parameter values and sample sizes derived as above for the most recent year for which data are available (2004). The model is then fitted to each artificial data set, the significance of the variation in year effects evaluated via a standard likelihood-ratio test, and the number of significant outcomes as a proportion of the total replicates adopted as an approximate measure of power. For a large number of models and species, repeated simulations are computer intensive thus the results below are based on 200 simulations for England, Scotland, Wales and Northern Ireland individually, but only 100 for Great Britain as a whole. These simulation sizes are certainly small enough to leave the resulting estimates of power (essentially binomial variables) prone to a degree of sampling error. For instance, it is readily shown by standard binomial theory (e.g. Freund & Walpole, 1987) that to estimate the proportion of significant outcomes from 200 trials as 0.5 is accompanied by a standard error of about 0.035. The standard error will reduce the more this proportion increases or decreases from 50%. This is small enough, therefore, for the estimates to function as a guide to the approximate size and nature of sample survey required to give an acceptable level of power, given that (as in any power analysis), even a precise estimate will inaccurately reflect the outcome of the field survey to a degree, as even the best estimates of the likely trend over time upon which simulations could be based will never mimic reality without some error. Experiments indicated that a batch of 500 simulations, which would for example reduce the standard error above only from 0.035 to 0.022, would require days of computing time for certain species and sampling structures,

3.2 The structure of the power study

Twelve species of significance in upland habitats were selected for detailed examination, and are listed in Table 1. We assume initially that the regular BBS will continue in its present form, at its current level, as a baseline against which to compare various forms of augmentation. For the simulation study, we need to set the number of these squares on which a species is present. To do this we take the mean site effect from Table 1, the best available indication of the current abundance of the species on upland BBS squares. If we then assume the number of birds on a square has a Poisson distribution with mean given by this value, we can readily deduce the probability of the bird being located in a given year. Since we also have, from Table 1, an estimate of the annual average numbers of squares recording the species in the BBS to date, these two quantities can then be used to estimate, crudely, the total number of squares at which we should regard the species as potentially present, and use this as the starting point for the simulation study, as in the example below. From there, each species is assumed to decline by 1.1% per year over a period of 25 years. This rate of decline amounts to a substantial loss over 25 years, approximately that required for the species to join the amber list of species of conservation concern. It is, then, the sort of decline we would very much hope to be able to detect.

Example

From Table 1, the mean site effect for Dipper is -0.9095 . A value of this magnitude implies a count of $\lambda = e^{-0.9095} = 0.4027$ birds at such a site. Further, since the probability mass function of the Poisson is given by:

$$
p(X=x; \lambda) = e^{-\lambda} \lambda^x / x! ,
$$

 $x = 0,1,2,3...$

the probability of such a site failing to register the species is given by $p = p(X=0;\lambda) = e^{-\lambda} = 0.6685$. That is, if the Poisson distribution is appropriate for each site at which the species is located there will be approximately two at which it is not recorded. Given an average from the BBS to date of 22 sites per year recording dippers, this implies using a value of about 66 sites per year (rounded to an integer) in the first year of the simulations. Beginning the simulated series with 66 occupied squares will therefore be expected to produce records of birds, on average over a large number of replicates, in around 22 squares in this initial year, mimicking the authentic BBS data. As we impose declines over time, this value will of course fall as the years progress, as will the mean number of birds per square. Conditional upon the parameter values, recorded presences or absences at a site are therefore independent between consecutive years. In reality, such a dependence might be expected to exist but for simplicity we ignore this detail here.

3.3 Augmented survey protocols

As power is not expected to be great for many upland species under the level of activity of the current BBS, we also consider three different means by which coverage of upland sites might be increased, and the extent to which power is improved as a consequence. We consider, for instance, a straight increase in the scale of the survey by adding 100%, or 300%, more squares, selected randomly in accordance with the BBS protocol. That is, simply increasing the scale of the BBS, retaining the sampling and field protocols, either two- or four-fold. Since this involves selection of squares at random, on average the number of squares initially containing the species will increase in proportion, the new squares having site effects drawn from the same normal distribution.

An increase of even 100% (i.e. a doubling of the total number), maintained over 25 consecutive years is a tall order if surveyors are to be voluntary recruits, or potentially expensive if the additional sites are surveyed on a professional basis. Two alternative means of, effectively, increasing the numbers of birds counted are also considered. Firstly, we assume existing volunteers are given the option of surveying one or more additional squares from those adjacent to their nominal square. It is assumed for the simulations that the bird density in the supplementary square is equal to that of the nominal square, as seems reasonable provided the extra square is selected independently of the number of birds present, e.g. it is a random adjoining square, or always the square due North, but surveyors are not permitted to choose the one most likely to be productive. Thus a simulated count is drawn from a Poisson distribution with mean two (or more, as appropriate) times that employed for the single square based on Table 1. It is assumed for this study that i) all of the volunteers respond to this request and continue the survey into a second square or ii) only half of all observers perform additional counts, but these survey two adjacent squares. It is assumed that whether a surveyor visits a second and third square is independent of the numbers of birds present; that is, surveyors with the most productive sites are not taken to be any more likely to take on a second square. Many upland squares are remote and considerable investment of time is required on behalf of the surveyor even to reach them and carry out a reasonably quick survey. For this reason it is believed many volunteers would be willing to extend the time spent surveying relative to their initial travelling time. An identical amount of land is therefore covered in both i) and ii). For model simulation, site effects are selected from the same normal distribution. For simplicity, we assume those sites in (ii) receiving enhanced effort are the same each year. For model fitting, the log of the number of squares surveyed (1, 2 or 3, effectively the 'area' of the site) is used as a model offset, which will be necessary in the more realistic, practical circumstances where there may be some change in the set of augmented sites, but is otherwise unnecessary as the change in numbers of birds with the increasing area covered can be accommodated by the estimation of the site effects. The offset has the advantage, even here, that site effects will reflect density (independent of area) and thus remain comparable with one another. Thus, the density (rather than the count) of birds at a site is set to be a (log-)linear function of the site and year effects only. Otherwise, model fitting proceeds similarly. Three is probably the maximum area that is likely to be feasible for a single observer.

A third means of enhancing the survey is considered as follows. The number of (randomly selected) squares is raised four-fold but the squares supplementary to the standard BBS are surveyed professionally not every year but only every 3 years on a regular basis. Thus in our 25 year artificial survey, the sites covered are doubled in years $1,4,7,10,13,16,19,22$ and 25. In a similar manner, we also considered a seven-fold increase in (randomly selected) squares every sixth year.

These basic augmentation strategies we shall denote by the letters (a-g) as follows:

- a) Continuation of the BBS in its current form
- b) Fourfold increase in the number of (random) squares covered $(= 4 x$ current survey effort)
- c) Twofold increase in the number of (random) squares covered $($ = doubling of current survey effort)
- d) Number of sites visited remains the same as the current BBS, but each surveyor also covers an additional random, adjacent square $($ = doubling of current survey effort)
- e) Number of sites visited remains the same as the current BBS, but 50% of surveyors each cover two additional adjacent squares. The remaining 50% survey a single, 1 km. square, in the traditional manner $($ = doubling of current survey effort)
- f) Continuation of the BBS in its current form, but with an additional 300% increase in squares every third year $($ = doubling of current survey effort)
- g) Continuation of the BBS in its current form, but with an additional 600% increase in squares every sixth year $($ = doubling of current survey effort)

Thus in (f) and (g) the standard BBS sites are surveyed annually, with three (six) times as many additional squares surveyed every third (sixth) year. These additional squares then are not surveyed annually, but the same squares are surveyed at every turn of the three- or six-year cycle. These schemes are investigated in turn matched to current sample sizes and parameter values for Great Britain, and also separately for England, Wales and Scotland (CEH landclass data do not exist for Northern Ireland).

Two additional analyses have been carried out, based solely upon the Great Britain data, which combine characteristics of more than one of the original set:

- h) A continuation of the BBS at current levels, plus an additional survey three times greater (in terms of sites) carried out every third year. In this additional survey, each surveyor covers 2 adjacent 1 km squares. The 'core' BBS data and the additional survey are then analysed together $(= 3 \times$ current survey effort)
- i) As above, but with each site in the additional survey extended to cover 3 adjacent 1 km. Squares ($= 4$ x current survey effort)

3.4 Approximating power for large numbers of species.

For a single species, large-scale simulation of any of the sampling strategies above is a time consuming exercise. Extension to the 50+ species recorded with any regularity on upland sites in this fashion is therefore computationally prohibitive. We therefore seek a more time-efficient means of approximating the power, with acceptable loss of accuracy.

To this end, we first turned to the various applications of strategy (a), a continuation of the current BBS, to each of the target species in each geographic area considered (England, Scotland, etc.) – a total of 49 species/region combinations of this kind are carried out in this report. The outcome of any set of simulations within the same sampling strategy is determined entirely by the number of sites assumed, and the parameter values (mean and variance of the estimated site effects), and is not otherwise connected to the species from which these are derived. Thus each power estimate can be treated as an independent observation and the set of estimates related via logistic regression to the variables from which it was derived. If the match between 'observed' power (from the simulations) and that 'predicted' (by regressing these on the parameters used in the simulation) is acceptable, this gives us a means of approximating power for large numbers of additional species, simply by applying the regression model to their own values for n, μ and σ .

We then used the estimates of power for the 13 target species and geographic areas considered for two of the augmented scenarios: (d) where the numbers of BBS sites remains the same, but where each surveyor also covers one additional adjacent square, and (f) where the current survey is augmented every three years by an additional three-fold increase in the numbers of random squares surveyed. Using these two additional logistic regression models, and species-specific parameters generated for the geographic regions, we then estimated the powers for all species occurring in uplands, for each geographic region, for these two scenarios.

4. RESULTS

4.1 The scale of Great Britain

The first set of analyses was carried out at the scale of Great Britain, because the CEH land-class categories used to define upland squares do not cover Northern Ireland (see above for the methods used to define upland in Northern Ireland). The standard Poisson regression models were fitted to each species to extract parameter values for use in the simulation study. For each species but two, the Pearson Chi-Squared goodness-of-fit statistic was between 1 and 2 times greater than its degrees of freedom, indicating no severe lack of fit. The exceptions, Golden Plover (4.37x d.f.) and Siskin (3.47x d.f.), both prone to flocking, suggested some overdispersion in the data with respect to the Poisson distribution. It should be borne in mind therefore that quoted figures for these two species may slightly overstate the power of the tests.

Power analyses matched to the BBS at the GB level reveal that for all but three of the 13 target species the existing survey, if continued at its present level, will detect the designated decline with >95% power. Since survey augmentation can only increase these values, further investigation is futile and in subsequent analyses at the GB scale we restrict attention to the remaining three species (Common Sandpiper, Dipper and Ring Ouzel). The current level of BBS coverage is estimated to detect the decline in these species with 89%, 59% and 56% power respectively.

The extent of the improvement in power under augmented schemes b-e are shown in Figure 1. For the Common Sandpiper all four result in close to 100% power. For the other species scheme (b) is (inevitably) better than (c) but while this difference is substantial for Ring Ouzel, even (c) provides $>$ 90% power for Dipper. Schemes (d) and (e), with volunteers visiting squares adjacent to their nominal sites, results are broadly comparable in magnitude to (c), though the total number of sites is rather less and requires only the retention of existing participants, and their willingness to extend the area they survey. [Note: Common Sandpiper (e) failed; likely to be ~100%] Schemes (f) and (g) raise the power to 81% and 84% respectively for the Ring Ouzel; for the other species each results in a figure $> 95\%$.

The two 'combined' strategies ('h' and 'i') above, with transient increases in numbers of sites and enlarged areas surveyed at those sites visited every third/sixth year, produce an almost total acceptance of a significant decline. Of the 12 species considered in the first stage of the analyses, it was only worth looking at the effects of three species for which the power of the current BBS coverage in Great Britain to detect declines was less than 90% and hence room for improvement. With three 1 km squares covered, all 100 simulations matched to the Common Sandpiper and Dipper produced statistically significant declines; 97 Ring Ouzel simulations similarly. Figures assuming only two squares at each site covered in the supplementary survey were only marginally less, the lowest figure of the three species being 95% (Ring Ouzel).

Figure 1 Results of simulations with current BBS (a) and augmented BBS sampling strategies (b) to (e) to estimate power to detect 25% declines for three species ($CS = \text{Common}$ Sandpiper, $DI = Dipper$ and $RZ = Ring$ Ouzel) in Great Britain. The Common Sandpiper simulation for strategy (e) failed (see text).

4.2 England

Once the scale of the analyses is reduced, matched to the constituent countries England, Wales and Scotland, the numbers of squares recording a species (and hence employed in the simulations) inevitably decrease. Thus, at the current level of coverage, six species (Common Sandpiper, Dipper, Ring Ouzel, Kestrel, Siskin and Tree Pipit) fall short of the 95% level of power that the latter three achieved on a GB scale. Once more we shall omit from further analyses those six species with associated power at this level under the present form of the BBS, along with Hooded Crow that is effectively absent from English BBS squares.

Scheme (b) is again the most beneficial adjustment; under a four-fold increase in the number of squares visited, only Common Sandpiper and Dipper still fall short of 90% power (Figure 2). In each case power is greater than, or approximately equal to, 80%. For these species and Ring Ouzel especially the gain under (b) is considerably greater than that under the existing regime, which gives power of only around 40% or less.

Figure 2 Results of simulations with current BBS (a) and augmented BBS sampling strategies (b) to (e) to estimate power to detect 25% declines for six species ($CS = \text{Common}$) Sandpiper, $DI = Dipper$, $K = Kestrel$, $RZ = Ring$ Ouzel, $SK = Siskin$ and $TP = Tree$ Pipit) in England.

4.3 Scotland

Scotland is the only country to record all of the thirteen species of interest in sufficient number for analysis. Of these, six species achieve 95% power at this scale; augmented analyses were carried out for the remaining seven (Common Sandpiper, Dipper, Kestrel, Ring Ouzel, Siskin, Snipe and Tree Pipit). Note that UK records of Hooded Crow are virtually restricted to Scotland, and Scottish results for this species are effectively interpretable also at the UK level.

Schemes (b-e) all produce power well over 90% for many species: Common Sandpiper, Hooded Crow, Red Grouse, Siskin and Snipe (Figure 3), some of which have Scotland as their British stronghold. At the opposite extreme, none of the investigated changes are predicted to raise power above even 25% for the rare Ring Ouzel. As with all previous analyses, schemes (c-e) tend to produce very similar results, none having a marked and consistent advantage.

Figure 3 Results of simulations with current BBS (a) and augmented BBS sampling strategies (b) to (e) to estimate power to detect $25%$ declines for nine species (CS = Common Sandpiper, $DI = Dipper$, $HC = Hooded Crow$, $K = Kestrel$, $RG = Red$ Grouse, $RZ =$ Ring Ouzel, $SK = S$ iskin, $SN = Snipe$, $TP = Tree Pipit$) in Scotland.

4.4 Wales

Golden Plover and Hooded Crow are effectively absent from Welsh BBS squares. Of the remaining species, only Song Thrush shows >95% power to detect the decline, thus augmented analyses were conducted for the remaining ten species.

Increasing the scale of the BBS even twofold provides almost 100% power for Pied Wagtail and Tree Pipit (Figure 4). The four-fold increase of scheme (b) is however required to raise power above 50% for species such as Kestrel and Snipe, but the relatively limited coverage in Wales means that even an increase of this scale leaves power below 40% for Common Sandpiper and Ring Ouzel**.**

4.5 Northern Ireland

In Northern Ireland, numbers of eight species were too low on upland squares to estimate parameters, or be included in the analyses. For the remaining seven species (Curlew, Dipper, Hooded Crow, Kestrel, Pied Wagtail, Snipe and Song Thrush), the power of the current BBS to detect a significant 25% decline was less than 95%, and hence simulations of augmented sampling were carried out for all seven.

For Dipper and Kestrel, none of the augmented scenarios, even quadrupling the upland sample annually, resulted in powers greater than 35% (see Figure 5). For Hooded Crow and Song Thrush (with powers of 94% and 84% respectively under the current sampling regime), all augmentations yielded very high power. For Curlew, some scenarios 'c' to 'g' resulted in estimates of power in excess of 80% but for Pied Wagtail and Snipe, the power achieved from those scenarios was between 56% and 78%.

Figure 5 Results of simulations with current BBS (a) and augmented BBS sampling strategies (b) to (g) to estimate power to detect 25% declines for seven species (CU = Curlew, $DI = Dipper, HC = Hooded Crow, K. = Kestrel, PW = Pied Wagtail, SN = Snipe, ST$ = Song Thrush) in Northern Ireland

4.6 Logistic approximations for large numbers of species

For the first step, we considered the 49 applications in this report of simulations based on scheme (a). We modelled the number of significant outcomes from the total replications (100 or 200) via logistic regression, as a function of n, μ and σ^2 , thus:

 $N_{sig} \sim Bin(N_{sim}, p)$ where $logit(p) = a_0 + a_1.n + a_2.\mu + a_3.\sigma^2$

where N_{sig} and N_{sim} are respectively the numbers of simulations with significant differences between the estimated year effects and the total number of simulations executed, and a_0 - a_3 are estimable coefficients.

Clearly this exercise introduces an additional level of approximation, but fitted values under this model show a reasonable fit to the simulated values (Figure 6). There is inevitably some scatter about the line of parity, but the model is certainly successful in predicting species of 'high' or 'low' power. We then fitted the Poisson model to BBS data for all species recorded in any number in upland regions, to obtain parameter values, and used the coefficients of the logistic model to predict power for these species without undergoing a full simulation exercise (Appendix).

Figure 6 Logistic regression model derived from the results of power estimates for the target species in scenario (a): a continuation of the BBS in its current form

This exercise was repeated for two alternative strategies ('d' and 'f'), obtaining new logistic regression coefficients that were used to estimate power analogously (shown in Figures 7 and 8).

Figure 8 Logistic regression model derived from the results of power estimates for the target species in scenario (f): a continuation of the BBS with an additional three-fold increase in the number of random squares every three years.

These results are listed in the Appendix and summarized in Table 2 below. Scenario a, the continuation of BBS in its current form, achieves $> 80\%$ power for 64% (79) of the species, while almost a half of species (48%) achieve this in England, 39% for Scotland and Wales and 25% for Northern Ireland. Scenario d, where each surveyor covers an additional random, adjacent square increased the amount of species achieving >80% power in all cases, with Wales, Scotland, England and GB all achieving this for 50% or more of the species. For Scenario f, the continuation of the BBS in its current form, but with an additional 300% increase in squares every third year, the percentage of species achieving >80% power is just slightly higher than that for scenario d.

Table 2 The number of different species falling into five different power categories from the results of scenarios a, d, and f for Great Britain, England, Scotland Wales and Northern Ireland. For full details see Appendix.

5. DISCUSSION & CONCLUSIONS

A number of the BBS augmentation scenarios simulated were selected to represent the same amount of survey effort – in terms of total number of BBS surveys carried out over time – albeit under different sampling regimes (including elements of non-annual surveys and enlarged sampling areas). In particular, scenarios 'c' to 'g' are all equivalent to a doubling of survey effort, per appropriate time period.

Despite the equivalent effort in carrying out the survey, the scenarios differ in other ways that influence the broad-scale efficiency. In scenario 'd', for example, surveyors could easily carry out the surveys on both adjacent squares during the same morning, requiring a single trip (actually two: one early and one later in the season) to the site, whereas scenario 'c' where the number of random squares is doubled would require close to twice the travelling time. Scenarios 'f' and 'g' address a different issue, related to deployment of volunteers versus professionals. Whereas it would be extremely difficult (impossible) to quadruple the number of volunteers every three years, it might be possible to hire fieldworkers every third year, with each fieldworker covering at least 30 sites over a three month period (assuming that only one random square in upland habitat could be surveyed per working day). The combination approaches in 'h' and 'i' could involve the use of professional fieldworkers every third year, but each carrying out surveys on two or more adjacent squares every day. In terms of total survey effort, scenario 'h' represents a three-fold increase and scenario 'i' a four-fold increase, over each three-year time period.

BBS surveys require two counting visits per season. Assuming a three month (60d) working field season, and a maximum of one 'random' square visit per day (hence 30 squares per fieldworker), and a maximum of three surveys on adjacent squares that could be carried out in a day, we calculated the following for Great Britain (with 300 upland squares). For the purposes of an initial assessment, we have assumed a cost of ca. £10k per three-month fieldworker position, including subsistence, administrative and mileage costs. Note that costs for the scenarios exclude overall volunteer recruitment, management and data inputting. The effort required to manage the collection of data from 300 volunteers would be about an eighth of the BBS operational budget. We start with an assumption that there are about 300 upland squares, currently surveyed every year, across Great Britain, which would be split between the constituent countries according to the proportion of upland that they contained.

Example 1. UK

Scenario a: Current state

300 random upland squares, carried out by approximately 300 volunteers (close to maximum volunteer capacity). Covered by existing BBS budget so additional cost ca **£0** over three years – but see above.

Scenario b: 4X BBS.

1200 random sites in total, 300 carried out by volunteers as above, and 900 extra squares, 30 each per season done by 30 professional fieldworkers, every year. Cost ca. £300k x 3y = **£900k** over three years.

Scenario c: 2X BBS.

600 random sites, carried out by approximately 300 volunteers as above, and 300 squares, 30 each by 10 fieldworkers, every year. Cost ca. £100 x 3y = **£300k** over three years.

Scenario d: Current BBS @ 2 x 1 km².

300 random 'doubled' sites, carried out by approximately 300 volunteers (needs piloting). Cost ca. **£0** over three years. **May be difficult to achieve**.

Scenario e: Current BBS + 0.5 @ 3 x1 km2

300 random sites (half 'tripled'), carried out by approximately 300 volunteers (needs piloting). Cost **£0**. over three years. **May be difficult to achieve**

Scenario f: BBS + 3 x BBS every 3rd Yr.

300 random sites, carried out by volunteers, plus 900 extra random sites (30 each by 30 professional fieldworkers) surveyed once every three years. Cost ca **£300k** over three years.

Scenario g: BBS + 6 x BBS every 6^{th} **Yr.**

300 random sites, carried out by 300 volunteers, plus 1800 extra random sites (30 each by 60 professional fieldworkers) surveyed every six years. Cost ca. **£600k** over six years (equivalent to **£300k** over three years).

Scenario h: BBS + 3 x BBS \circledcirc 2 x1 km² every 3rd Yr.

300 random sites, carried out by volunteers, plus 900 extra sites – each a double square, surveyed by 30 professionals (each doing 30 'sites'), once every three years. Cost **£300k** over three years.

Scenario i: BBS + 3 x BBS \circledcirc 3x 1km² every 3rd Yr:

300 random sites, carried out by volunteers, plus 900 extra sites – each a triple square, surveyed by 30 professionals (each doing 30 'sites'), once every three years. Cost **£300k** over three years.

Hence, at a first look, for a number of these scenarios, assuming roughly equal upland coverage across the four countries, the additional costs of deploying professionals to increase the power of the BBS to detect declines in upland bird populations could be, on average, ca. £75k over three years, an average of £25k per year per country. Again, note that this includes only additional fieldworker costs. In fact, the breakdown of the number of upland BBS squares currently surveyed per year, on average, in the UK and its constituent countries is: UK 343, England 132 (38% of total), Scotland 125 (36% of total), Wales 61 (18% of total) and Northern Ireland (9%).

To conclude, although scheme b, that quadrupled survey effort, was predictably the scheme that resulted in the largest increase in power, other options also provided good increases in power, most notably the two options f and g, involving a 3-fold and 6-fold increase in coverage every third or sixth year respectively, although this scheme would require additional funding to employ fieldworkers. Other than for a and b, there was in fact relatively little difference between the increase in power provided by the different options. However, scenarios involving increasing effort through additional adjacent squares rather than additional random squares would require far less effort from fieldworkers in terms of travel time so are more efficient.

Acknowledgements

We are very grateful to Sian Whitehead and Andrew Owens from the Countryside Council for Wales, Neil McCullagh from the Environment Heritage Service in Northern Ireland, Mark Eaton and David Gibbons from RSPB, and particularly Andy Brown and Allan Drewitt from Natural England, for obtaining the financial support for this work from their respective organisations. We would also like to thank Jeremy Greenwood, Stephen Baillie and Sarah Eglington for their expert advice and valuable contributions to this report, Stuart Newson for his help in categorising and extracting the BBS data on which the analyses are based, members of the Upland Bird Monitoring Group for comments received following the presentation of results, and Mandy T Andrews for formatting the final version of the report.

References

Freeman, S.N., Noble, D.G., Newson, S.E. & Baillie, S.R. (2007) Modelling population changes using data from different surveys: the Common Birds Census and the Breeding Bird Survey. *Bird Study*, **54**, 61-72.

Freund, J.E. & Walpole, R.E. (1987) *Mathematical Statistics*, 4th edn. Prentice-Hall.

Haines-Young, R.H., Barr, C.J., Black, H.I.J., Briggs, D.J., Bunce, R.G.H., Clarke, R.T. Cooper, A. Dawson, F.H., Firbank, L.G., Fuller, R.M., Furse, M.T., Gillespie, M.K., Hill, R., Hornung, M., Howard, D.C., McCann, T., Morecroft, M.D., Petit, S., Sier, A.R.J., Smart, S.M., Smith, G.M., Stott, A.P., Stuart, R.C. & Watkins, J.W. (2000) *Accounting for nature: assessing habitats in the UK countryside*. DETR, London ISBN 1 85112 460 8.

APPENDIX

Table A1 Logistic regression-based estimates of power (p) for all common breeding species recorded on upland sites in Great Britain, based on scenario (a).

Table A5 Logistic regression-based estimates of power (p) for all common breeding species recorded on upland sites in Great Britain, based on scenario (d).

Table A9 Logistic regression-based estimates of power (p) for all common breeding species recorded on upland sites in Great Britain, based on scenario (f).

Table A10 Logistic regression-based estimates of power (p) for all common breeding species recorded on upland sites in England, based on scenario (f).

 \mathbf{r}

Table A13 Logistic regression-based estimates of power (p) for all common breeding species recorded on upland sites in Northern Ireland, based on scenario (a).

Table A15 Logistic regression-based estimates of power (p) for all common breeding species recorded on upland sites in Northern Ireland, based on scenario (f).