A study to develop the scope for monitoring landscape-scale biodiversity impacts of agri-environment schemes in England

J.T. Staley¹, G.M. Siriwardena², S.M. Smart¹, R.S. O’Connor¹, I.G. Henderson², S.K. Jarvis¹, N. Jones³, S.N. Freeman¹, J.W. Redhead¹, C. Carvell¹, C. Hallam³, M. Jitlal¹

¹ NERC Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire, OX10 8BB
² British Trust for Ornithology, The Nunnery, Thetford, Norfolk, IP24 2PU
³ FERA Science Ltd, National Agri-food Innovation Campus, Sand Hutton, York, YO41 1LZ

Natural England project number: ECM 42922

Natural England project officer: Susanna Phillips

3.2 Scoring of taxa evidence ................................................................. 62
3.3 Contrasting local and landscape AES gradients ............................. 65
  3.31 Potential National Character Areas for proposed landscape monitoring ........ 66
  3.32 Existing monitoring schemes’ coverage of the local and landscape AES gradients ..... 71
3.4 Designing a landscape monitoring scheme: conclusions .................... 71
4 Species monitoring strategy (work-packages 3 & 5) ............................................ 73
  4.1 Designing a species monitoring approach: protocols to survey individual taxa ........ 73
  4.1.1 Monitoring pollinating insects and butterflies ................................................. 73
  4.1.2 Monitoring moths ....................................................................................... 80
  4.1.3 Monitoring birds ......................................................................................... 83
  4.1.4 Monitoring terrestrial mammals ................................................................. 86
  4.1.5 Monitoring bats ......................................................................................... 89
  4.1.6 Monitoring amphibians and reptiles ........................................................... 91
  4.1.7 Monitoring plants ....................................................................................... 95
4.2 Assessments of AES implementation ...................................................... 98
4.3 Estimates of number of sampling units needed including power analyses ........ 99
  4.4.1 The power of monitoring to detect effects of AES interventions .................. 99
  4.4.2 Estimating the effects of AES on wildlife counts: a power analysis .......... 100
4.4 Estimates of the resources needed to monitor landscape-scale effects of AES ...... 105
4.5 Species monitoring strategy: conclusions .................................................. 112
5 Review of analytical techniques for landscape-scale biodiversity data ............ 113
  5.1 Summary of strengths & weaknesses of modelling methods for landscape-scale detection of AES effects ................................................................. 113
  5.1.1 Stratified random sampling ................................................................. 113
  5.1.2 Bayesian Belief Networks (BBN) ....................................................... 113
  5.1.3 Habitat Suitability modelling and use of existing models ...................... 114
  5.1.4 Remote sensing .................................................................................... 114
  5.1.5 Occupancy Detection models ............................................................... 115
  5.1.6 Generalised Linear Models and Mixed Models (GLM and GLMM) ........... 116
  5.1.7 Structural Equation Models (SEM) ...................................................... 117
  5.2 Conclusions from summary of strengths and weaknesses of analytical techniques .... 118
  5.3 Detailed review of analytical techniques for landscape-scale biodiversity data .. 122
    5.3.1 Generalised Linear Models (GLM) / Generalised Linear Mixed Models (GLMM) .. 122
    5.3.2 Structural equation models ................................................................. 126
    5.3.3 Bayesian belief networks ................................................................. 128
5.3.4 Informative priors ....................................................................................................... 131
5.3.5 Incorporating existing datasets ................................................................................ 132
5.3.6 Habitat suitability mapping ...................................................................................... 138
5.3.7 Remote sensing ......................................................................................................... 142
5.3.8 Stratified random sampling ...................................................................................... 146
5.3.9 Cumulative landscape effects .................................................................................. 148
5.4 General issues identified by analytical scoping ........................................................... 152
  5.4.1 Taxonomic level of interest ...................................................................................... 152
  5.4.2 Rarity ...................................................................................................................... 152
  5.4.3 Legacy effects ........................................................................................................ 152
  5.4.4 Spatial vs. temporal analysis .................................................................................. 153
  5.4.5 Extension to non-target taxa .................................................................................. 153
  5.4.6 New types of data ................................................................................................... 153
5.5 Consideration of analytical power ............................................................................... 154
5.6 Analytical approaches in relation to different survey designs ...................................... 154
5.7 Detailed review of analytical techniques: conclusion and comments ...................... 155
Executive summary

Agri-environment schemes (AES) are the most significant mechanism to deliver environmental policy within England, and include the new Countryside Stewardship (CS) AES launched in 2016. Previous studies of AES have found mixed evidence for effects on biodiversity, and have largely focussed on responses of wildlife taxa within individual AES options or agreements, with a few assessments of effects at the landscape or national scale where possible. Here, we present the findings of a study to scope approaches to monitoring biodiversity responses to AES at a landscape scale in England, specifically considering impacts beyond farm or agreement boundaries.

The objectives of this study were to:

1) Review the evidence for key species groups, focussing on landscape-scale studies where available, to collate information on (i) the strength and type of evidence for responses to AES (ii) field survey techniques, (iii) habitat and landscape variables shown to influence the response of taxa to AES interventions and (iv) existing monitoring schemes that could contribute to landscape-scale AES monitoring.

2) Consult stakeholders on: (i) which species groups to include in the evidence review, (ii) the relevant spatial and temporal scales for monitoring a range of taxa, (iii) appropriate field survey techniques, (iv) the use of volunteers for AES monitoring and (iv) broader comments around landscape-scale monitoring of AES effects.

3) Develop a score for level of AES intervention on potential study units within National Character Areas (NCAs), and test the use of NCAs as landscape units for landscape-scale monitoring.

4) Design a species monitoring strategy, including detailed field survey protocols, a framework to assess AES implementation, estimates of the replication required and indicative costs for landscape-scale monitoring of AES effects.

5) Scope analytical approaches to quantify the relationship between the extent of AES intervention and the responses of taxa at local and landscape scales, and including AES implementation success and key habitat variables. Consider the possible role of predictive modelling to scale up beyond those landscape areas monitored for responses to AES.

Approach. A detailed, semi-systematic review of published literature, reports and expert species reviews was carried out to meet Objective 1 above, and was also informed by expert scoring exercises and information collated during a stakeholder workshop (Objective 2). Information collated during the review was scored in relation to the strength and type of evidence. The scores were combined with current Environmental Stewardship uptake data to calculate evidence-based gradients of AES intervention. The potential to use a study design based on crossed gradients at local (1 × 1km) and landscape (3 × 3km) scale AES intervention was explored, which included identifying which NCAs could support this design, and the characteristics of these NCAs (e.g. cover of major land use types and species pools, Objective 3). Field survey protocols were
designed for nine species groups, in consultation with taxon experts, together with a framework for rapid scoring of AES implementation. A generic power analysis was conducted, based on simulations of changes to wildlife populations over time at sites under varying extent of AES intervention. The power of a landscape monitoring scheme was estimated in relation to variables including the number of sites, number of years and the mean and variance of expected initial abundance (which varies with the response variable and taxa/species monitored). Modular cost estimates were derived for field surveys across the full range of taxa for which protocols were designed, and indicative costs are presented for a monitoring project to assess several potential research questions over different time-scales (Objective 4). Potential analytical approaches were scoped through a detailed review, based both on studies identified during the evidence review and on new approaches not yet applied to AES research. A summary of the strengths and weaknesses of seven analytical approaches in the context of landscape-scale monitoring of AES impacts is presented, together with a range of model schema for testing landscape-scale AES effects on species abundance (Objective 5).

Key findings

Background and summary of evidence from review and stakeholder workshop

- The type and strength of evidence relating to AES effects differed with species group, as did the interpretation of how ‘landscape-scale’ effects of AES were assessed. The most frequent approach was to quantify ‘background’ landscape variables (e.g. landscape heterogeneity), to test whether these correlated with local-scale effects of AES management. This interpretation was used across the broadest range of taxa reviewed, including for pollinating insects (bees and hoverflies), butterflies, moths, birds and plants. Whilst this approach usefully informs decisions regarding the landscape contexts in which AES effort may be best placed to maximise local-scale benefits, it does not demonstrate impacts of AES that extend beyond land under AES management, or are sustained over time.

- For mobile organisms, there is a need to demonstrate that AES benefits extend beyond changes to localised abundance, which could be due to short-term redistribution of organisms in response to increased resources supplied by AES management. Existing monitoring data have been used to assess changes in populations of bird and butterfly species, in 1km squares, in relation to uptake of AES schemes/option groups across spatial scales of 1-25km. Spill-over from land under AES management onto adjacent non-AES land has recently been assessed in two studies for pollinating insects. We thus interpret ‘landscape-scale’ impacts of AES as requiring a test of whether AES benefits (i) are sustained over time in terms of population growth (or reduced rates of population decrease), or (ii) extend more widely onto surrounding land not under AES management (spill-over effects), or ideally both.

- During the stakeholder workshop birds, pollinating insects, butterflies and moths were ranked as important candidates for both landscape-scale and population monitoring. Monitoring at the landscape-scale was ranked slightly more highly than population change for some plant groups. Terrestrial mammals were ranked low for both landscape-scale and population monitoring, with the exception of brown hares. For bats, the importance of landscape-scale monitoring was scored more highly than monitoring of population change. The importance of
landscape-scale monitoring was scored lower for amphibians and some reptile species, probably due to these being less mobile species.

- Evidence for the effects of AES management on rare, conservation-priority (section 41) invertebrates was found less often than for more widespread, common species of butterflies, moths and pollinators. Workshop participants commented that rare ‘bespoke’ section 41 invertebrates may be better monitored through targeted surveys at specific sites (e.g. SSSIs), than with a landscape-scale AES monitoring scheme designed to cover a range of species groups.

**Design and testing of AES gradients**

- Evidence from the review of AES benefits across a range of option groups was strongest for birds (divided into boundary and in-field functional groups), pollinating insects and butterflies, for which evidence-based AES uptake gradients were calculated. The AES gradients overlapped between these taxa in the majority of NCAs, with the exception of nine NCAs in which no major agricultural land category dominated, so an average AES gradient was calculated. Sampling sites for these four taxa can thus be co-located, potentially saving resources and facilitating the identification of co-benefits.

- Within individual NCAs there was little evidence that the AES gradients varied consistently with background landscape characteristics, in terms of proportional cover of arable or semi-natural grassland, or landscape diversity, so measurement of AES effects should not be confounded by such variation. This is probably due to NCAs being defined to encompass land with similar characteristics and lends support to the concept of using NCAs as a focus for monitoring. At larger spatial scales AES gradients may co-vary more strongly with background landscape characteristics.

- Seventeen NCAs were identified with the potential to support landscape sampling units along full contrasting local (1km square) and landscape (3×3km) AES gradients, which would allow the local and landscape effects of AES interventions to be assessed. These 17 NCAs were larger than the average NCA, but did not differ substantially in terms of coverage of major land cover categories, and included NCAs dominated by each of arable and intensive grassland. Using a crossed gradient design, nine or sixteen 3×3km landscape sampling units could be surveyed in an NCA.

**Field monitoring proposals and analytical approaches**

- Field survey protocols were designed to survey pollinating insects, butterflies, moths, birds, terrestrial mammals, bats, amphibians and reptiles. These protocols were designed to be compatible with some existing volunteer monitoring scheme data, to allow their use for contextual data as discussed below. Surveying entire landscape sampling units would enable spill-over of mobile organisms between adjacent 1km squares with high and low or no AES intervention to be assessed, in addition to surveying just the central 1km square which would provide an assessment of changes in populations, in relation to local and landscape AES intervention. Where possible, protocols survey more than one taxon on the same visit (e.g.
butterflies and other pollinating insects), but for many taxa there is a trade-off between optimal survey design and the capacity to combine surveys with other taxa.

- Existing monitoring schemes, such as the Breeding Bird Survey and Wider Countryside Butterfly Survey, can provide contextual data on national population abundance and trends, and for scaling up of results from targeted monitoring. We do not advocate assessing landscape-scale impacts of the new CS scheme based solely on existing monitoring schemes, as: existing sites are not located to maximise the contrast in AES uptake, high turn-over of sites means that data cannot be guaranteed over a particular 5-year period, and scheme survey methods may not be intensive enough to quantify local-scale populations accurately in order to detect potential spill-over of AES benefits. Moreover, the Wild Pollinator and Farm Wildlife package of the new CS scheme was designed partly to meet the conservation requirements of pollinating insects, for which no monitoring scheme currently exists that includes solitary bees and hoverflies.

- Survey protocol design and resource estimates were based on employing professional field surveyors. Participants at the stakeholder workshop commented that existing volunteers are often less motivated to undertake additional data collection tasks, to collect data from prescribed locations, or to survey intensively farmed sites where rare species are seldom recorded and a proportion of the surveying may consist of recording zero abundance. The recent National Pollinator and Pollination Monitoring Framework project also considered the potential to use volunteers in the context of designing a new pollinator monitoring scheme, and concluded that taxonomic expertise was a limiting factor in the expansion of pollinator monitoring by volunteers, leading to a trade-off between cost and the quality of data if volunteers are used.

- The power of a monitoring scheme to detect AES effects on population changes varies strongly with the response variable and taxa monitored, represented by differences in the initial mean count per site. An estimated power of 0.8 may be achieved through monitoring for a minimum of 5 years at 100 – 200 sites for response variables where initial counts are greater than 10 per site, for example the abundance of common pollinator, butterfly or bird species, or aggregated abundance across multiple species. For species with initial mean counts of less than 10, monitoring over longer time scales might be needed to reach a similar level of power.

- The modular format of both the protocols and resource estimates for the field surveys allow any combination of taxa to be included in the monitoring programme, depending on policy priorities. Similarly, the indicative costs for several entire monitoring schemes scenarios, summarised along with the research question addressed and pros and cons of each scenario, illustrate alternative approaches depending on which aspect of ‘landscape monitoring’ is prioritised for AES monitoring. There may be opportunities for cost savings if compromises are made as to the ideal sampling framework for particular taxa, facilitating the combination of sampling for different groups into single survey visits. Reduced replication (sample size) could also reduce cost, with consequences for likely statistical power, but using volunteer survey effort would not be feasible.
• The use of independent local and landscape AES gradients varying from highest to lowest uptake would strengthen the ability to quantify relationships between AES uptake, quality of implementation and changes in species abundance, and allow the construction of testable hypotheses. Predictive model performance is likely to be improved by the inclusion of other covariates known to influence biological responses to changing management, for example habitat variables or the legacy of previous AES schemes. However, high collinearity would mean that shared variance cannot be uniquely attributed to either of the correlated variables, so their effects could not be separated. There will be a limit to the number of covariates that can be included.

• Generalised linear (mixed) models have been used extensively in AES research, and are also likely to be the basis of any analytical framework. Structural Equation Models are a promising approach to investigate direct and indirect relationships between AES uptake and taxon abundance. A Bayesian framework may allow a combination of these and other analytical approaches (using independent data sets) into a single model. The addition of other methods such as habitat suitability mapping, and the inclusion of existing databases, may allow extension of models to include prediction outside of surveyed sites.

• Remote sensing and existing recording schemes can provide data that could be used to support conclusions from field survey data, provide valuable contextual information, and to enable prediction of responses outside of surveyed areas. However, due to the temporal and spatial resolution of remote sensing and recording scheme data, and the methods used, they are not suitable to assess landscape-scale effects of AES without additional field survey data. Methods relying on prior information or expert opinion (e.g. Bayesian Belief Networks) were considered to have low ability to detect or predict landscape-scale effects of AES, and could lead to inaccurate conclusions, so are not recommended.

Summary. The evidence review highlighted that interpretations of landscape effects in the context of AES differ. We conclude landscape scale monitoring should include a test of whether AES benefits (i) are sustained over time in terms of population change, or (ii) extend more widely onto surrounding land not under AES management (spill-over effects), or ideally both. Which interpretation is chosen will determine the conclusions can be drawn from the results of a landscape AES monitoring scheme, and thus which policy questions can be informed. A proposal for a modular, flexible, detailed field monitoring scheme has been developed to allow monitoring of the effects of landscape and local scale AES interventions for a wide range of taxa, using a novel study design based on independent, evidence-based AES gradients at local and landscape scales.
1 Introduction

Agri-environment schemes (AES) represent the most significant environmental policy delivery mechanism within England, both in terms of expenditure and scale of coverage of the countryside. The underlying aim of AES in England (and throughout the UK) has been to foster positive management of environmental features, including wildlife, public access, landscapes and the historic environment. Monitoring of AES has provided mixed evidence for overall scheme success, especially with respect to effects on biodiversity (Kleijn et al. 2011; Oliver 2014).

1.1 Why monitor biodiversity responses to agri-environment schemes at the landscape scale?

The majority of research into the effects of AES interventions on wildlife taxa has been focussed on monitoring of individual AES options, or AES agreements (e.g. Mountford et al. 2013). This scale of research is necessary for testing and improving AES management prescriptions (e.g. Blake et al. 2010; Staley et al. 2016), and to determine whether AES interventions are benefitting target taxa on land directly under AES management, which is perhaps the most relevant spatial scale for less mobile taxa such as plants (Walker et al. 2006; Marshall 2009). However, for mobile taxa such as birds and some invertebrates (e.g. pollinating insects or butterflies), there is the potential for organisms to move onto AES land when increased resources are available, without this having a sustained effect on populations over time or across the wider countryside. Recognition of this possible ‘honeypot’ effect (resulting from the spatial redistribution of individuals across the landscape) has led to some tests of whether responses to AES interventions are more sustained at the population level and from local to landscape scales, but only for a limited range of taxa (e.g. Brereton et al. 2007; Baker et al. 2012; Carvell et al. 2015). More rigorous assessments of individual AES, such as the new Countryside Stewardship (CS) scheme launched in England in 2016, across larger spatial scales and longer timescales are needed to determine whether schemes are having a sustained positive effect across a range of policy-relevant mobile taxa. Moreover, a principal aim of CS is to maximize the cost-effectiveness of AES spending by focusing a restricted budget on higher value areas, with the expectation that the management in these areas will lead to an export of environmental benefits to the wider landscape. It is critical, therefore, that the success of this export process is monitored, as well as responses to management undertaken within AES farms.

Landscape-scale effects of AES interventions have been previously interpreted in several ways, including: (1) testing effects of landscape characteristics (such as complexity) on local-scale responses of taxa to AES interventions, (2) assessing effects of AES coverage in the wider landscape on short-term, localised responses of taxa, (3) assessing the effects of existing and previous AES coverage in the wider landscape on long-term population responses, for example through the use of existing monitoring scheme data for birds and butterflies, usually within 1km squares, and (4) assessing the relationship between abundance of mobile taxa on AES land and nearby areas not under AES management, to test for potential ‘spill-over’ effects. Here, we interpret the assessment of landscape-scale AES effects to require assessments of whether effects are both sustained over time, through effects on rates of change in abundance or population sizes.
(3 above), and to be demonstrated across larger spatial scales than land directly managed under AES (4 above). Such effects could, in principle, be assessed via analyses of data from ongoing surveys or via new, bespoke monitoring effort, or via a combination of the two. A more detailed discussion of the interpretation of landscape-scale effects, together with taxon-specific evidence for the different interpretations, is given in Section 2.

To scope possibilities for landscape-scale monitoring of the new CS AES in England, a comprehensive, semi-systematic evidence review was conducted across a range of taxa, with input from taxon experts through a stakeholder workshop (Section 2). The potential to use AES uptake data to select possible sampling units along evidence-based AES gradients at varying spatial scales was explored, together with the potential for National Character Areas within England (the landscape units used for targeting of the new CS scheme) to support such a monitoring scheme design, as units with consistent background landscape characteristics (Section 3). Detailed monitoring methodologies were developed, including: protocols to sample a large number of species from a range of taxa, a framework for rapid assessment of AES implementation, consideration of the number of sampling units required to provide adequate power, and indicative resource estimates for a number of potential monitoring scheme scenarios (Section 4). A rigorous review of the possible statistical approaches that could be used to analyse monitoring scheme data, together with scoping the potential contribution of existing datasets such as those obtained through remote sensing or ongoing, existing monitoring, is presented in Section 5.
2 Review of evidence and expert opinion for taxa responses to AES management and suitability for monitoring (work-package 2)

2.1 Literature review of evidence

To assess the evidence for the effects of AES interventions on defined taxonomic groups, semi-systematic literature searches were conducted of peer-reviewed research, grey literature and documented expert opinion. Several species groups were reviewed: six that were identified as likely candidates for landscape monitoring by Natural England (birds, butterflies, pollinating insects, section 41 invertebrates\(^1\), bats, arable plants), and additional taxa proposed during a stakeholder workshop (terrestrial mammals, amphibians, reptiles, moths; Section 2.2).

The evidence was collated and summarized for each taxon, in terms of: (i) the amount/strength of evidence for AES effects within different management option groups, (ii) the type of evidence (e.g. short term abundance response vs population change, local vs. landscape scale responses, expert opinion vs. published research results) and (iii) the response variables assessed. Evidence collated regarding existing monitoring schemes, other sources of relevant data and survey / monitoring techniques were used to inform work on the species monitoring strategy (Section 4). Details of the literature review methodology and data collation are in Appendix A, and the evidence is summarized for each taxon below (Sections 2.3 and 2.4).

2.2 Workshop to consult stakeholder and expert opinion

A stakeholder workshop was jointly organised by BTO and CEH in February 2016. Representatives from twelve external organisations were invited, selected to complement the taxon knowledge within CEH and BTO. The workshop was attended by nineteen representatives from nine external organisations and an independent entomologist, in addition to NE, Defra, JNCC and the project team (Appendix B for details). Five workshop sessions were run, two of which involved scoring taxa (i) in relation to their priority for landscape monitoring using criteria such as policy relevance, ease of sampling and value as a proxy for other taxa (session B categories and scores in Appendix B) and (ii) in terms of the spatial and temporal scales that are necessary and feasible for monitoring (session C1). Two sessions captured information on (i) key evidence sources of taxa responses to AES management (session A), to supplement the evidence review (Section 2.1), and (ii) details of survey techniques (session C2). All four of these sessions used small break-out groups for scoring or to capture evidence efficiently from workshop participants. A final session (C3) engaged all participants in a general discussion about the potential design of a landscape-scale monitoring scheme. In addition, unstructured comments were gathered from participants about broader issues relating to AES monitoring. Details of the workshop programme, participants, scores, discussion notes from two of the sessions and unstructured stakeholder comments received during and after the workshop are provided in Appendix B.

---

\(^1\) Species defined as being ‘of principal importance’ for conservation under section 41 of the Natural Environment and Communities (NERC) Act 2006
Scoring of species groups during session B confirmed that workshop participants thought the six broad groups (birds, bats, butterflies, pollinating invertebrates, section 41 invertebrates and arable plants) chosen as a focus for the evidence review broadly met the objectives of this study, in terms of factors such as policy relevance and responsiveness to AES management (Table B, Appendix B). Other species groups that scored highly in session B included amphibians, reptiles, mammals (other than bats) and moths, which were added to the evidence review (Section 2.3). Evidence captured during session A on responsiveness to AES management is incorporated into the taxon literature review (Section 2.1), and evidence gathered in session C2 on survey techniques is used in the design of the species monitoring strategy (Section 4).

Scores for the importance, feasibility and necessity of monitoring at varying temporal and spatial scales indicate that participants thought monitoring of effects on population change (i.e. over temporal scales greater than 2-3 years or single generations), and at a landscape scale (in addition to local scales), were important for birds, pollinating insects and butterflies. Landscape monitoring was ranked slightly more highly than population change for some plant groups (Appendix B, see section C1 for details of all the scoring results). Habitat quality indices were considered sufficient for some common and arable plants, and measures of spatial variation in abundance and cover were ranked highly by some workshop participants for arable plants, grassland fungi and indicator plant groups. The necessity for landscape scale monitoring and measures of population change in relation to AES were given low scores for some terrestrial mammal species, with the exception of brown hares. Population change effects were ranked low for groups of bat species, and a comment was made that regional bat population trends are already monitored; landscape-scale monitoring importance was scored highly for all bat species groups due to their high mobility. The importance of landscape scale monitoring was scored low for amphibians and some reptile species, but more highly for four widespread reptile species. The importance of landscape scale and population change monitoring also scored highly for moths (Section C1).

The final, open discussion session (C3) focused on possible designs of modules for a landscape monitoring scheme. In addition to taxon-specific comments, several participants highlighted that asking volunteer recorders to collect additional data results in a large reduction in the number of volunteers (Appendix B session C3 notes). It was also agreed that, although sampling multiple species groups at the same time is desirable, in reality optimal sampling methods and timings often differ and require different taxonomic expertise, so the potential to sample multiple taxa with one protocol without compromising survey effectiveness is limited to a few specific cases. Written comments included the suggestion that the spatial scale of monitoring may need to differ between ‘mosaic’ species (as defined in Natural England’s CS targeting) and species with ‘bespoke’ habitat requirements. This was echoed by another participant’s comment suggesting that rare/localized species monitoring may be focused on SSSIs, while widespread species are more likely to be detected by landscape-level AES monitoring.

The focus for the plant evidence review was discussed during the workshop. Some participants felt the focus on rare arable plant species was too narrow, while others suggested plants may not be good candidates for landscape-scale monitoring due to their lack of mobility. It was suggested that the review of plants for monitoring be broadened to look at AES management effects on early successional species as a general group, to extend the review beyond arable habitats. It was
also agreed that the assessment of plant communities could also have a role in landscape-scale monitoring in terms of quantifying habitat quality for more mobile species.

### 2.3 Evidence reviews for species groups

#### 2.3.1 Pollinating insects

For the purposes of this review, pollinating insects comprise all bee and hoverfly species, of which there are around 268 and 284 species respectively (Carvell et al. 2016). Three of the studies reviewed also included butterflies in their definition of pollinating insects (Pywell et al. 2011b; Korpela et al. 2013; Scheper et al. 2013). A total of 65 research papers or studies in the grey literature addressing pollinating insects were reviewed. For two section 41 species of bumblebee (*Bombus humilis* and *B. sylvarum*), recommendations for AES management under the Glastir Monitoring and Evaluation Programme (GMEP) based on ecological literature and expert opinion were also recorded. Twenty-five studies were specific to the UK, 22 in England, two in Scotland and one in Wales. Thirty-four studies came from other single European countries, whilst five studies used data from multiple European studies and one used a global dataset.

A large number of studies within the non-AES bracket consider organic farming (Figure 1), representing a form of reduced input agriculture predominantly focusing on crops, but also commonly with significant implications for the whole farming system, such as more mixed farming and denser hedgerows. The remaining option groupings are represented by few studies or none and so further discussion will focus on field margin treatments, low input grassland, general AES and organic farming.

Fifteen of the reviewed studies presented direct tests of AES with a counterfactual control (e.g. wildflower margin vs. crop edge), and a further four presented direct tests of AES without a counterfactual control (e.g. comparison of several AES options; Figure 1). Thirty of the studies reviewed presented a test of habitat(s) relevant to AES with a counterfactual control and twelve presented results that tested the effects of an AES-relevant habitat without a counterfactual control.
Figure 1 Summary of the number of studies which tested broad AES option grouping for pollinating insects, and the proportion in these groups covered by different types of studies. AES CF = direct test of AES option against a counterfactual control, which provided a ‘non-AES agricultural management as usual’ comparison, (e.g. AES grass buffer strip vs crop edge as counterfactual). AES NCF = direct test of AES using a non-counterfactual control (e.g. comparing AES options only against each other, or variants of management within the same option). NAES CF = test of habitat(s) relevant to AES or management using a counterfactual control; NAES NCF = test of an AES relevant habitat or management without a counterfactual control.

2.3.1.1 Metrics and taxonomic coverage of pollinator studies

Fifty-five of the studies on pollinating insects presented analyses using metrics of total abundance within defined taxonomic groups (e.g. all pollinating insects, all bees, bumblebees, hoverflies), 47 presented analyses of species richness and six used measures of diversity. Data were aggregated at varying taxonomic resolutions and far more studies presented results for all bees or all bumblebees than for hoverflies or solitary bees (Table 1). This taxonomic bias towards bumblebees is more pronounced among studies that reported abundance for individual species; across the 16 studies that do so, 14 species of bumblebee are represented, compared to four species of hoverfly (with results from just two studies) and one species of solitary bee, *Osmia bicolor*, (from a single study). Within the bumblebee species, three section 41 species are represented (*Bombus muscorum*, *B. ruderatus* and *B. ruderarius*).
<table>
<thead>
<tr>
<th>Grouping measured</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>All pollinating insects</td>
<td>2</td>
</tr>
<tr>
<td>Bees, Lepidoptera and hoverflies</td>
<td>1</td>
</tr>
<tr>
<td>Bees and hoverflies</td>
<td>1</td>
</tr>
<tr>
<td>Bumblebees and butterflies</td>
<td>2</td>
</tr>
<tr>
<td>All bees</td>
<td>19</td>
</tr>
<tr>
<td>All wild bees</td>
<td>5</td>
</tr>
<tr>
<td>Bumblebees and honeybees</td>
<td>1</td>
</tr>
<tr>
<td>Bumblebees</td>
<td>29</td>
</tr>
<tr>
<td>Hoverflies</td>
<td>13</td>
</tr>
<tr>
<td>Solitary bees</td>
<td>9</td>
</tr>
<tr>
<td>Red list bees</td>
<td>2</td>
</tr>
</tbody>
</table>

**Table 1** The broad taxonomic classifications used in studies of AES impacts on pollinating insects (only some studies also reported abundance at species level and/or species richness).

### 2.3.1.2 Spatial and temporal scale of pollinator studies

All of the studies addressing insect pollinators were classed as spatial because all utilised treatment groups separated in space and considered the effects of intervention over short time periods; 38 were conducted over one year, 13 over two years, eight over three years and two were over four years.

### 2.3.1.3 Evidence for response of pollinators to AES management from the research literature

#### 2.3.1.3.1 Field margins and floral resource based interventions

Several AES option groups are based on improving habitat quality in margins surrounding fields, or providing patches of floral resources, and these are the option types studied most often for impacts on pollinating insects (Figure 1). Margins supplemented with sown wildflower seed mixes, or pollinator specific mixes, were found almost always to have a positive effect on the local abundance and species richness of pollinating insects, particularly when an AES option was tested specifically (Figures 2 and 3).

Several studies of bumblebees compared multiple options not only against a crop control, but also against each other, in general showing that whilst margins that had not been sown with a flower seed mix (grass buffer strips, uncropped margins and conservation headlands) had greater pollinator abundance or species richness than crop edges, sown flower margins most often outperform these other options (e.g. Carvell *et al.* 2007). Further, many studies included a measure of flower abundance, and some showed that significant effects of wildflower margins are due to increased floral resources (Pywell *et al.* 2005; Marshall *et al.* 2006; Scheper *et al.* 2015).
Figure 2 Summary of the response of measures of total abundance of insect pollinators to AES option groups, under the different study type classifications. AES CF = direct test of AES using a counterfactual control; AES NCF = direct test of AES using a non-counterfactual control; NAES CF = test of AES relevant habitat using a counterfactual control; NAES NCF = test of an AES relevant habitat without a counterfactual control. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$. Green bars show the number records for recommendations within specific option groups across all available species based informed expert opinion resources.

Two studies directly tested spill-over effects of wildflower strips on pollinators, i.e. whether AES interventions can increase pollinator abundance on land that is not under AES management, but is in the vicinity of a wildflower strip. Carvell et al. (2015) measured the impacts of wildflower patches of different sizes on both local bumblebee densities and densities in the landscape surrounding the patch. Local (patch-level) densities of males and queens of three out of four bumblebee species were significantly increased on wildflower patches compared to controls. Within the surrounding landscape (based on transects within 1km of the focal patch), bumblebee density was decreased at small (0.25 ha) wildflower patches, suggesting that positive local impacts were at least partly due to the redistribution of individuals. However, the effect of the
size of wildflower patches was significant, as around the large (1 ha) wildflower patches bumblebee density was not reduced by AES intervention. Furthermore, bumblebee densities and responses to wildflower patches at both scales were influenced by landscape context (determined by the proportion of arable land, Carvell et al. 2011).

Figure 3 Summary of the response of pollinating insect species richness to AES option groups, under the different study type classifications. AES CF = direct test of AES using a counterfactual control; AES NCF = direct test of AES using a non-counterfactual control; NAES CF = test of AES relevant habitat using a counterfactual control; NAES NCF = test of an AES relevant habitat without counterfactual control. A non-significant effect was recorded when a result was not significant at p ≤ 0.05, but was at p ≤ 0.1.

Jönsson et al. (2015) found the presence of wild flower strips and the degree of landscape heterogeneity (determined by the proportion of semi-natural grassland and uncultivated field borders) influenced the abundance of bumblebees, solitary bees and hoverflies differently at the landscape / farm scale. Abundance was analysed at the whole farm scale, with data pooled across transects along field borders that were either adjacent, or up to 800m from wild flower strips. Bumblebee abundance was greater on farms with flower strips present than without, regardless of landscape heterogeneity, while solitary bee abundance did not differ significantly with either variable. For hoverflies a more complex picture emerged, where by the presence of wildflower
strips only increased abundance across the whole farm when landscape heterogeneity was low. Jönsson et al. (2015) also showed the distance over which local AES interventions affect abundance differed between pollinator taxa: bumblebee abundance on farms with wildflower strips was higher on both adjacent field borders and those further away (up to 800m); there was no significant difference in abundance for solitary bees between wildflower strips and any field borders (although numbers were low), while for hoverflies abundance was greater on wildflower strips than adjacent field borders, but not those further away. These differences are probably linked to the both the movement capacity of these different groups (solitary bees and hoverflies move over smaller distances than bumblebees) and their foraging behaviour; lacking a nest site to return to, hoverflies will generally forage over shorter distances than bees and will be also be searching for mating and egg laying habitats.

2.3.1.3.2 Low-input grassland

The results classed as low-input grassland under AES come from two studies, based on the same experiment addressing the impacts of AES meadows in Switzerland (Albrecht et al. 2007a; Albrecht et al. 2007b). Using insect visitation to phytometer plants, Albrecht et al. (2007a) showed that intensively managed farmland adjacent to Ecological Compensation Area (ECA) meadows had significantly higher abundance and species richness of pollinators as a whole group, bumblebees, solitary bees and hoverflies compared to fields adjacent to intensively managed grassland. Trap-nesting solitary bees exhibited higher species richness but not abundance (Albrecht et al. 2007b). Albrecht et al. (2007a) also found increased pollination services in ECA vs. intensively managed grassland. None of the four studies that addressed low input grassland with counterfactuals in a non-AES context found a significant, positive effect on abundance, whereas three (of four) studies found evidence for a significant effect on species richness (Figure 3).

2.3.1.3.3 Organic farming

Studies addressing organic agriculture dominate those not directly addressing an AES scheme or option, and focused on comparisons of reduced agricultural inputs such as pesticides, herbicides and fertiliser with conventional management in arable habitats. Overall, organic farming methods significantly increased both abundance and species richness of pollinating insects, although again there is a taxonomic bias; three studies reported results for hoverflies as a distinct group as opposed to twelve focusing on bees.

2.3.1.3.4 Species-specific responses of pollinators to AES management

A small number of studies also reported results for the abundance of individual species, the majority of which are bumblebees (Figure 4). The species level results for hoverflies are from two studies on the effects of flower-rich margins, once considering multiple species (Haenke et al. 2009) and another focusing on Episyrphus balteatus (the Marmalade fly, Macleod 1999). The result for Osmia bicornis (the Red mason bee) comes from a study comparing organic and conventional fields (Holzschuh et al. 2010). It is of note that the abundance of Bombus lapidarius, B. pascuorum and the B. terrestris / lucorum complex generally responded positively to all margin types, with the exception of conservation headlands, which consistently had no effect (Figure 4).
Wilkinson et al. (2012) is one of two studies which explicitly address the impacts on insect pollinators of targeted AES management for a species of another taxon, the Corncrake (*Crex crex*). The promotion of early and late tall grassland vegetation cover had significant positive effects on the abundance of *Bombus lucorum*, *B. muscorum* (a section 41 species), *B. hortorum* and *B. distinguendus* – with only *B. jonellus* showing no response. While rare species can clearly be attracted to well-established flower margins in some regions, given appropriate forage plant cover (Carvell et al. 2007), those species for which significant abundance responses are most commonly observed are also the more common species and so a lack of significant effect for other species (e.g. *Bombus hortorum*, *B. pratorum* and *B. ruderatus*) is probably due in part to insufficient sample sizes, given their relative rarity or lower activity at the time of sampling (Carvell et al. 2007; Kuussaari et al. 2011). The fact that the number of results reported is lower for these species is also indicative that they were not detected in sufficient numbers for statistical tests to be conducted.

2.3.1.3.5 Trait based responses of pollinators to AES management

Beyond broad taxonomic groups, some studies have split pollinator taxa into trait-based groups to explore whether this influences responses to AES. Four studies have divided bumblebees into short-tongued and long-tongued species, all concerned with responses to different field margin options, and three specifically assessing options implemented under ELS or old Countryside...
Stewardship (old CS) prescriptions (Pywell et al. 2005; Öckinger & Smith 2006; Pywell et al. 2006; Heard et al. 2007). In the UK-based studies, both groups showed positive responses to all options tested, apart from conservation headlands, mirroring the results for all bumblebees combined.

One study differentiated hoverfly species based on whether their larvae were aphidophagous or detritivorous (Kohler et al. 2007), but this distinction was only made when exploring the effects of distance to nature reserves on abundance. The abundance of both guilds was significantly greater in fields in closer proximity to semi-natural habitats, but for species richness there was only a significant positive effect for aphidophagous species, which is at least indicative that different responses can be detected for different functional groups.

2.3.1.4 Evidence for response of pollinators to AES management from expert species review

Evidence based on expert judgement was found for two pollinating insect species, the bumblebees *Bombus humulis* (Brown banded carder bee) and *B. sylvarum* (Shrill carder bee). Their favoured habitat is species-rich grassland or meadows, and as such grassland management prescriptions dominate recommendations, in contrast to the results from empirical studies which focus on arable field margins (Figure 2). These are both rare, section 41 species, so are less likely to have been recorded from field studies.

2.3.1.5 The influence of landscape on responses of pollinating invertebrates to AES

Empirical studies on pollinators divide into those showing an influence of landscape characteristics on local scale abundance responses to AES and those showing an effect of localised AES management on landscape scale responses; only three studies reviewed addressed the latter. As discussed above (Section 2.3.1.3.1), Carvell et al. (2015) found local interventions had a negative effect on the abundance of queen and male bumblebees in the wider landscape if AES interventions were small, but no effect if large wildflower patches were sown. Bumblebee abundance was increased on AES wildflower patches partly due to redistribution of individuals from the wider landscape onto the smaller sized patches (Carvell et al. 2015), but findings point to overall population increases where larger flower patches are introduced, demonstrating the need to monitor abundance of spatially-linked populations in order to interpret spill-over effects (point 4 in interpretations of landscape effects described in Section 1.1 above).

In an analysis that combined data from Pywell et al. (2006) and Carvell et al. (2007), Pywell et al. (2012) found that species richness of rare bumblebee species was significantly higher on farms with pollinator flower and nectar supplements, and that this local scale effect was linked with species richness of rare bumblebee species in the wider landscape (10km × 10km). Jönsson et al. (2015) found the response of pollinator abundance at the landscape (whole farm) scale to presence of a wildflower strip on the farm and landscape heterogeneity varied between taxa (Section 2.3.1.3.1). This emphasises the importance of considering the differing spatial scales at which different taxonomic groupings and species are likely to respond to AES, and their autecology, when interpreting results. Further, it shows problems may arise from considering the response of aggregated abundance variables across groups or species, as more abundant taxa may mask the response, or lack of response, of other species.
Two studies of wildflower-based AES interventions show the amount of AES in the landscape surrounding farms can have beneficial local effects, and modify the response to local management. Pywell et al. (2006) showed that the number of AES agreements with pollinator and wildflower interventions in a 10 km x 10 km area surrounding farms had a significant, positive effect on the number of long-tongued bumblebees recorded. In an study of four European countries Scheper et al. (2015) found increased early floral resource availability in the surrounding landscape increased the number of bumblebee individuals and species found on wildflower strips, but both metrics decreased on control crops. For solitary bees, in contrast, the amount of late floral resources in the landscape had a significant, positive effect of the number of individuals recorded on control strips but decreased the number found on wild flower strips, and there was no effect on species richness (Scheper et al. 2015). The extent of AES implementation in the surrounding landscape can thus have an important influence on effectiveness at the local scale, but this interaction may vary with taxonomic group.

Several studies have assessed the effect of landscape ‘type’, classifying landscapes based on the amount and / or diversity of semi-natural habitats found within them. Four of these studies fall into the general AES category, two of which were meta-analyses, and one utilised data collected for both arable and pastoral systems in six European countries (Batary et al. 2011; Concepción et al. 2012; Scheper et al. 2013). They cover an array of different AES options / schemes across continents and so effects can only be attributed to general AES intervention, but their sample sizes and geographic coverage are large. All four studies show significant, positive effects of AES on the abundance and species richness of pollinating insects, indicating that AES interventions broadly have positive impacts. Central to the results of all four studies is the finding that AES effectiveness is increased in less complex, more agriculturally homogeneous, landscapes. In highly intensively farmed landscapes (e.g. <1% semi-natural habitats) effectiveness was also low, but such landscapes are less relevant to the UK. Concepción et al. (2012) suggest that, based on this relationship, AES interventions are most effective in landscapes of intermediate complexity. From their meta-analysis, Scheper et al. (2013) found the magnitude of the effect of AES interventions was larger for cropland than grassland systems. Kennedy et al. (2013), using a global dataset from 39 studies, did not directly address AES, but instead the impacts of organic farming and local habitat diversity (based on whether fields had a mix of cropped and non-cropped habitat including field margins and hedgerows). Diverse fields had significantly higher abundance of both wild social bees and solitary bees, and significantly higher species richness of solitary bees and organic farming increased abundance and species richness significantly for all taxonomic groups tested (all bees, wild social bees and solitary bees).

Several other studies assessing landscape effects on local responses to specific AES options or management have found that along with positive effects of AES, there were also significant increases in abundance and species richness on sites surrounded by a higher proportion of semi-natural habitats (e.g. Holzschuh et al. 2008; Holzschuh et al. 2010), whilst others did not (e.g. Brittain et al. 2010; Wood et al. 2015a). Several studies report higher abundance of bumblebees on floral margins / plots when there was a higher proportion of agricultural land in the surrounding landscape (Heard et al. 2007; Carvell et al. 2011) or differing effects for bumblebees, solitary bees and hoverflies (Holzschuh et al. 2008; Jönsson et al. 2015).
Two studies assessed how the response of AES may be dependent on landscape context but were not be included in Figures 1-4 as they assessed how landscape modified the effects of AES rather than effects of AES per se. Concepción et al. (2008) explored whether differences in bee abundance and species richness between farms under AES and farms without AES were affected by the diversity of habitats within a landscape, the level of connectivity and the complexity local field shapes; none of these factors were found to have a significant effect on either abundance or species richness. Grass et al. (2016) assessed the abundance and species richness of hoverflies and bees on wildflower plots in relation to the proportion of arable land and presence of other wildflower plots in the surrounding landscape within a 500 m radius. Results showed a negative effect of arable land on bee abundance and bee species richness, whereas hoverfly abundance and species richness were positively affected by the presence of other wildflower plots, suggesting that an increase in wildflower interventions in the surrounding landscape facilitates higher local abundance.

Together, these findings suggest that surrounding landscapes can have important consequences for the effectiveness of AES interventions; interventions are likely to have the largest effects in landscapes with a low proportion of semi-natural habitat (<20%), but in severely cleared landscapes (<1% semi-natural habitat) positive effects will be lessened (Scheper et al. 2013). However, beyond the studies that have taken a large scale experimental or meta-analysis approach (Batary et al. 2011; Concepción et al. 2012; Kennedy et al. 2013; Scheper et al. 2013), variation in the definition of the scale of landscape and the landscape characteristics measured make it difficult to draw specific conclusions about the effect of landscape characteristics on local scale abundance.

2.3.1.6 Is there evidence that AES management addresses limiting factors for pollinators?

Bee populations may be limited by three factors: food resources, nesting resources and refugia from incidental risks (e.g. pesticide exposure, Dicks et al. 2015). Strong evidence exists for a limiting role of floral food resources for wild bee populations, in contrast to the other two factors (Roulston & Goodell 2011; Dicks et al. 2015), although pollinator declines may be due to multiple drivers and also differ between bee species or functional groups (Vanbergen & Initiative 2013). Dicks et al. (2015) suggest that many AES options available within ES could provide food, nesting habitats and refugia for pollinators. The floral units required per ha farmland per month were estimated for six dominant, crop-pollinating, wild bee species, and compared to the floral resources potentially provided by AES options (Dicks et al. 2015). A critical gap in floral resources for bees in March and April was identified, before many sown AES wildflower options are flowering. Early flowering woody hedgerow species such as blackthorn may partially fill this gap (Dicks et al. 2015), but this will partly depend on hedgerow management which is also targeted by AES (Staley et al. 2012). Thus, there is strong evidence that floral resources may limit populations of some wild bee species, especially in early spring months. The lack of evidence for effects of nesting resources and refugia may reflect a focus on floral resources in the AES literature rather than proof that these are not limiting for some wild bee species; for example, some ground-nesting solitary bee species may be limited by a lack of bare ground for nesting (Potts et al. 2005).
There is no equivalent assessment of limiting resources for hoverflies within an AES context, but Meyer et al. (2009) emphasise that hoverflies switch between habitats for foraging, mating, overwintering and larval feeding. Hoverfly larval feeding habits are variable, ranging from aphidophagous species to saprophagous and phytophagous species. Hoverflies as a group thus require a diversity of micro- and macro-habitats, including woody habitats for shelter and overwintering, as well as floral resources for adults to feed on (Meyer et al. 2009).

2.3.1.8 Gaps in the evidence of AES management effects on pollinators

Most of the evidence reviewed for pollinating insects has focused on options targeted at increasing floral resource provision, low input grasslands, or the beneficial impacts of general AES rather than option- or habitat-specific actions. Floral resources can be an important limiting factor for pollinating insects (Section 2.3.1.6) and some studies have shown floral resource availability was the mechanism through which abundance or species richness was increased. However, for some species AES interventions may be acting on other limiting factors. For example, there is little or no direct evidence for AES effectiveness at providing nesting habitats, although several option types could provide these (e.g. grassland management, hedgerow management, stonefaced hedges, stone wall management; Dicks et al. 2015). By measuring the number of cells occupied by pupae in traps nests for solitary bees, some studies have shown that interventions also create habitats preferential for nesting – but this could also be a measure of redistribution. One study has tested the effects of options on colony reproductive success, using the buff-tailed bumblebee (Bombus terrestris), showing no difference between colonies located on farms with conservation headlands compared to those which without (Goulson et al. 2002). Just one study found an increase in abundance of nest searching queen bumblebees associated with hedgerows, species rich grassland and grass margins / beetle banks in Scotland (Lye et al. 2009). Evidence is needed for the mechanisms through which other AES management can support pollinating insects, and also to determine which resources limit different pollinator taxa (bumblebees vs. solitary bees vs. hoverflies) functional groups and species. Further, there is little research available that addresses the impacts of targeted management for other taxonomic groups on insect pollinators, with the exception of Wilkinson et al. (2012). Finally, although Wood et al. (2015b) was able to detect higher nest densities of two bumblebee species on HLS farms with wildflower margins than ELS farms, constituting larger populations on these enhanced farms within a single year, no study on pollinators has tested the effects of AES interventions on changes in populations over time, in contrast to research on birds and butterflies reviewed below.

2.3.2 Butterflies

A total of 52 studies addressing butterflies were reviewed: 25 were specific to the UK (24 in England, one in Wales), 27 came from other European countries and one utilised data from England, France, Germany and Italy. Of the studies reviewed, 36 presented direct tests of AES with a counterfactual control and nine presented direct tests of AES without a counterfactual control. Twenty-two of the studies reviewed presented a test of an AES-relevant habitat with a counterfactual control and eight presented results which tested the effects of an AES relevant habitat without a counterfactual control.
Similar to pollinating insects, the dominant AES option groupings assessed for butterflies are field margin treatments, organic farming, low-input grassland and more general grassland management options (Figure 5). As with pollinating insects, studies in the organic grouping generally represent reduced input arable agriculture, whereas low-input grassland options represent reduced-input pastoral systems, with studies on the latter that directly tested AES options mainly assessing the Environmental Compensation Area (ECA) AES in Switzerland.

Figure 5 Summary of the number of studies which tested broad AES option groupings for butterflies and the proportion in these groups covered by different study types. AES CF = direct test of AES management using a counterfactual control; AES NCF = direct test of AES using a non-counterfactual control; NAES CF = test of AES relevant habitat using a counterfactual control; NAES NCF = test of an AES relevant habitat without a counterfactual control.

2.3.2.1 Metrics and taxonomic coverage of butterfly studies

Forty of the studies reviewed for butterflies present aggregated abundance measures, whilst 31 test species richness. Nineteen studies assessed the abundance of individual species, with a total of 47 species covered (42 of which are found in the UK; Figure 6).

Most studies considered all butterflies as an aggregate group, although four studies combined butterflies and moths. Two studies classified species into groups of mobile and sedentary species, three considered specialist and generalist species as discrete groups and two studies focused specifically of species in the genus *Thymelicus*. Three studies (Brereton 2002, 2005; Oliver 2014) have used data from the UK Butterfly Monitoring Scheme (UKBMS) to explore the
response of butterflies to AES, classified on habitat specialisms and UK Biodiversity Action Plan (BAP) classifications.

Figure 6 The number of studies addressing the abundance of 42 individual butterfly species as response variables.

2.3.2.2 Spatial and temporal scale of butterfly studies

Forty-six of the studies addressing butterflies were classed as spatial and six classed as temporal; the latter used time series data to assess the change over time in relation to AES intervention.

2.3.2.3 Evidence for response of butterflies to AES management from the research literature

2.3.2.3.1 Landscape-scale responses of butterflies to AES schemes

The impacts of specific AES schemes – old Countryside Stewardship (old CS), Entry Level Stewardship (ELS) and Environmentally Sensitive Areas (ESAs) - dominate tests for effects on aggregated butterfly abundance (Figure 7). Most of these results are from three studies that utilised the UK Butterfly Monitoring Scheme data to test for the effects of AES uptake on abundance, both across butterflies and several different sub-groupings (Brereton 2002, 2005; Oliver 2014). Brereton (2005) used the same dataset as (Brereton 2002), but with an additional two years of data, both comparing trends in abundance for butterflies at sites covered by old Countryside Stewardship (old CS) agreements with those that were not, but (Brereton 2002) also presented data for individual species (Section 3.5.1.4). Considering the aggregated abundance of all (39) butterfly species (Brereton 2005), there was no significant difference in trends between old CS and non-AES sites, and this result was repeated when only specialist (19) species were considered. When six species of grassland specialist butterfly requiring a mosaic of grassland and scrub habitats were considered it was found that declines were significantly larger on sites covered under old CS, which was also the case when considering six species classed as farmland.
indicators. However, when considering only BAP priority butterfly species it was found that trends were significantly more positive on sites with old CS implemented than those without. For short turf specialists there was a non-significant positive effect of old CS implementation.

**Figure 7** Summary of the response of measures of total abundance of butterflies to AES option groups, under the different study type classifications. AES CF = direct test of AES using a counterfactual control; AES NCF = direct test of AES using a non-counterfactual control; NAES CF = test of AES relevant habitat using a counterfactual control; NAES NCF = test of an AES relevant habitat without a counterfactual control. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$. Green bars show the number of recommendations within specific option groups across all available species based summaries of expert opinion.

Oliver (2014) also utilised the UKBMS dataset, analysing data from 2006 – 2011 to investigate the impacts of old Countryside Stewardship, ESA and ELS & HLS implementation on population trends, using the proportion of a land covered by AES agreements within 1km and 3km buffers surrounding each sampling site. A similar analysis investigated the relationship between mean abundance recorded at Wider Countryside Butterfly Scheme (WCBS) sites in 2010, and extent of AES intervention. For the UKBMS analyses, the only significant relationship found was for the proportion of land covered by old CS options thought to be beneficial to butterflies within a 3km buffer around sites. Analyses considering WCBS sites found significant, positive, relationships between abundance and both total ESA cover in a 1km buffers and old CS cover in a 3km buffer. The cover of ELS and old CS options beneficial to butterflies both had significant, positive effects on butterfly abundance at a 3km resolution.
Figure 8 Summary of the response of butterfly species richness to AES option groups, under the different study type classifications. AES CF = direct test of AES using a counterfactual control; AES NCF = direct test of AES using a non-counterfactual control; NAES CF = test of AES relevant habitat using a counterfactual control; NAES NCF = test of an AES relevant habitat without a counterfactual control. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$.

2.3.2.3.2 Field margins and wildflower supplements

Eight studies reviewed examined the effect of AES margin treatments against counterfactual controls (Figure 7). Four of these provided tests of wildflower margins or pollinator flower mixes, with three finding significant positive effects with regard to adult abundance (Meek et al. 2002; Potts et al. 2009; Korpela et al. 2013) and one finding no effect (Brereton 2005; Figure 7). However, Potts et al. (2009) found that the addition of a wildflower mix to low input grassland had a negative effect on butterfly larval abundance. Three studies addressed species richness, two of which reported significant positive effects (Meek et al. 2002; Aviron et al. 2009; Potts et al. 2009; Figure 8). Evidence is mixed for the impacts of grass buffer strips, with two of the four studies addressing AES options within this group finding significant, positive effects compared to controls.

Of the four studies to address wildflower margins, pollinator mixes and grass margins not directly linked to an AES option, there were positive effects on both species richness and abundance, although positive effects were not significant for one study (Alanen et al. 2011).
2.3.2.3.3 Low-input grassland and extensification

Across studies considering low input grassland management and extensification, effects were generally positive for both butterfly abundance and species richness. Of note are three studies addressing low-input grassland under agri-environment with counterfactual controls come from studies of ECA meadows, under the Swiss AES. All three studies found significantly higher butterfly species richness in ECA meadows (Aviron et al. 2007; Roth et al. 2008; Aviron et al. 2009). Further, two found that management regimes that left parts of ECA meadows unmown also served to increase butterfly species richness (Cizek et al. 2011; Kühne et al. 2015).

2.3.2.3.4 Organic farming

Ten studies assessed the impacts of organic farming on butterflies as an aggregated group, one of which only presented results for butterflies and moths combined (De Snoo et al. 1998). In six studies, land under organic management was found to have significantly higher numbers of butterflies, including Feber et al. (1997), which showed that the abundance of non-pest species was higher but there was no effect on the abundance of the pest species Pieris brassicae and P. rapae. One study reported no effect of organic farming on butterfly abundance, and another found no effect on the abundance of meadow specialist butterfly species. Ekroos et al. (2008) found that, whilst farms with a larger area of field boundary habitat supported a higher abundance and species richness of butterflies, there was no significant difference between organic and conventional farming. Three of the ten studies found organic farming had significantly greater butterfly species richness (De Snoo et al. 1998; Rundlöf & Smith 2006; Rundlöf et al. 2008) and four no effect on species richness (Weibull et al. 2000; Feber et al. 2007; Ekroos et al. 2008; Brittain et al. 2010).

2.3.2.3.5 Evidence for response of individual species

Brereton (2002) compared the trends in species over time on UKBMS sites with old CS AES options to those without. Significant effects were seen for many species, some of which increased in response to old CS, e.g. Dark green fritillary (Argynnis aglaja) and Silver-studded blue (Plebeius argus), whilst others were negative (e.g. Brimstone Gonepteryx rhamni and Small blue Cupido minimus). On calcareous grassland sites, significant effects of the old CS scheme were found for just two of 20 species: positive for the Chalkhill blue (Polyommatus coridon) and negative for the Wall brown (Lasiommata megera). Brereton (2005) also compared trends for High brown fritillary (Argynnis adippe) and Marsh fritillary (Euphydryas aurinia), both section 41 species, on sites with Biodiversity Action Plan habitats covered by old CS agreements and those not. The mean population decline for the Marsh fritillary was significantly lower for BAP sites covered by old CS than those that were not, and whilst statistical tests could not be used for the High brown fritillary, those sites under old CS had stable populations and those without showed a mean decline.
Figure 9 Summary of the test responses for 42 individual butterfly species across AES option grouping and the significance of effects. Species with an * are classified as section 41 under the NERC Act in England. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$.

Oliver (2014) examined whether AES coverage in the surrounding landscape affected abundance of individual species on WCBMS sites, with a significant effect of AES found for 19 species. For ELS coverage, six species showed positive responses, e.g. the Meadow brown (*Maniola jurtina*), the Small copper (*Lycaena phlaeas*) and the Ringlet (*Aphantopus hyperantus*) were positively affected by both coverage of ELS in general and coverage of butterfly beneficial options. For ESA cover, there were significant positive responses to both general and specific option coverage (e.g. the Wall brown, Meadow brown and Small tortoiseshell *Aglais urticae*), whilst others only responded to total cover or cover of options considered beneficial to butterflies. Positive responses were most numerous with regard to old CS coverage, with six species showing positive response to total old CS coverage general and butterfly specific options and seven species to only butterfly specific options, while one (the Green-veined white) was only responsive to coverage of all CS options (Figure 9).

Oliver (2014) also tested whether there were changes in temporal trends in abundance for individual species after the implementation of AES management. For each species, models were run for the data from all UKBMS sites for which AES had been implemented to see whether a linear model or broken stick model provided a better fit to the data – with broken stick models indicating a change in trends after the implementation of AES in either a 1km or 3km buffer surrounding a site. Evidence was found of a change in temporal trend following introduction of AES for eight of the 21 species considered. All of the evidence for a response to AES came from models considering old CS and ESA agreements, perhaps as these were the longer established schemes.

At the local scale, three studies on field margins assessed Meadow brown abundance, one showing that grass margins enhanced with wildflowers supported higher Meadow brown abundance than crop margins (Meek *et al.* 2002), and two studies showing that field borders with
6m and 2m grass margins also had higher abundance than borders with no grass margins (Field et al. 2005, 2006). The Small heath and Ringlet were significantly more abundant on field margins supplemented with wildflower seed mix (Meek et al. 2002); however no effect was recorded for twelve other species.

Two studies tested the impacts of organic farming on individual butterfly species using counterfactual controls. De Snoo et al. (1998) assessed whether reduced pesticide and herbicide application affected the abundance of butterfly species on potato and wheat field margins. There were significant responses for on organic margins for both crops for five of six species, with increased floral resources suggested as the driver. Feber et al. (2007) found that five of twelve species tested showed significantly higher abundances on organic vs conventional farms.

Two grey literature studies tested the effect of AES interventions on a section 41 butterfly species. Ellis et al. (2012) found evidence for a positive response to landscape scale management by the Marsh fritillary butterfly in two locations in the UK. Following implementation of HLS management in Dartmoor National Park (Devon, UK), larval web abundance increased (1,082%) and patch occupancy increased from six patches in 2005 to 18 in 2010. After implementation of AES management in the South Wessex Downs ESA (Dorset, UK), there was a significant increase in abundance of the Marsh fritillary in the region (a 278% increase over 20 years) over a 10 year period, compared to a significant decline across the rest of the UK (Ellis et al. 2012). Staley et al. (2014) found that hedgerows managed under the ELS option EB1 (cut once every two years in early autumn) did not have significantly higher numbers of Brown hairstreak eggs than a control hedge, but one variant of the HLS and ELS option EB3 (cut at most once every three years in early autumn) had significantly greater numbers of eggs than control hedges.

2.3.2.3.6 Trait-based responses of butterflies to AES management

The majority of studies have addressed either the abundance or species richness of all butterflies, but a few functional groupings have also been considered. Three studies distinguished specialist from generalist species, based on larval host plant use and habitat specialism (Brereton 2005; Aviron et al. 2011; Korpela et al. 2013). Two showed specialist butterflies increased in abundance on floral margins, although Aviron et al. (2011) found no effect on specialist species richness and Brereton (2005) showed no difference between trends in abundance for specialist species between sites covered by old CS and those not. Two studies grouped species based on adult mobility, a category linked to specialism, in that specialist species often exhibit low mobility. Pywell et al. (2011a) found that complex wildflower field margins significantly increased the abundance of mobile species, but not sedentary species.

2.3.2.4 Evidence for the response of butterflies to AES management from expert species review

Information based on informed expert opinion was collated for nine species of butterfly (all classified as section 41) from Natural England CS species factsheets. For three fritillary species (High brown fritillary, Marsh fritillary and Pearl bordered fritillary), evidence was also available from the Glastir species reviews. A broader range of AES management options are deemed important (as with pollinating insects) than those addressed by empirical studies (Figure 7), with
a particular emphasis on options for grassland management, as many of these section 41 species are grassland specialists.

2.3.2.5 The influence of landscape on local responses of butterflies to AES

Some studies address whether agricultural practices in the wider landscape can affect the impacts of AES at a local scale. Aviron et al. (2011) showed that species richness of generalist butterflies in fields was positively related to the percentage cover of wildflower strips found within a 200m radius. However, this was not the case for specialist species, which were influenced by the number of plant species found within a 100m² in the field. Gabriel et al. (2010) showed that farms in organic hotspots (landscapes with a greater abundance of organic farms) had a higher abundance of butterflies, and Hodgson et al. (2010) similarly showed that the amount of organic farming in a landscape had a positive effect on butterfly abundance, but the magnitude of this effect was much smaller than for local (farm-scale) organic farming. However, Aviron et al. (2009) found that the proportion of the surrounding landscape (in a 200m radius) either covered under ECA or under crop did not affect butterfly species richness on meadows or arable fields. Rundlöf & Smith (2006) and Rundlöf et al. (2007) showed butterfly abundance and species richness increased at local scales on organic farms located in homogenous agricultural landscapes, but not those in more heterogeneous landscapes.

Several studies tested whether the character of the landscape surrounding sites altered local-scale abundance of butterflies, providing mixed evidence. Using a measure of landscape diversity based on coverage of different agricultural, semi-natural and built environment land cover types, Weibull et al. (2000) found that the abundance of butterflies was positively affected by landscape heterogeneity at a large scale (5km × 5km grid surrounding a farm), whereas species richness and diversity were only responsive at the 400m × 400m scale. Several studies have focused on the effects of woodland habitats in the surrounding landscapes as many butterfly species have a requirement for shelter, with mixed results. Korpela et al. (2013) found that the proportion of forest surrounding farms had a significant influence on the number of specialist butterfly species found on wildflower flower strips and whilst Mazalova et al. (2015) found no difference between less intensive grassland management and controls, the abundance of butterflies on a site was positively related to the amount of tree belts in the surrounding landscape. However, Ekroos et al. (2008) found that distance from woodland had no effect on abundance or richness of butterflies on conventional and organic farms.

Together, these sources of evidence suggest that the local benefits of AES management for butterflies are highest in a more intensively farmed landscapes and, complementary to this, that increasing the amount of positive management in the surrounding landscape will support higher local abundance on farms. However, as with pollinating insects, the different approaches and scales used to quantify landscape characteristics will influence results, and this is further complicated by the fact that different species are likely to respond to AES at different scales, making generalised conclusions difficult.
2.3.2.6 *Is there evidence that AES management addresses limiting factors for butterflies?*

Several studies suggest that butterfly responses to AES management are due to the provision of limiting resources, though none explicitly demonstrate this. Similar to pollinating insects, many empirical studies show that increases in adult butterfly abundance or species richness relate to local increases in floral resources, but whether this a limiting resource for farm-scale (or larger scale) populations is not addressed. An important limiting resource for many butterfly species is known to be the density of their larval host plant, and more specifically their larval host plant occurring in suitable conditions (Thomas 1983; Thomas *et al.* 1986; Warren 1987; Thomas 1991; Bourn & Thomas 1992; Morris *et al.* 1993; Thomas *et al.* 2001; Botham *et al.* 2011; Merckx *et al.* 2012). The improved population trends of the Silver spotted skipper (*Hesperia comma*) and the Chalkhill blue on sites under AES management has been attributed to grazing regimes increasing the availability of their host plants found in suitable microhabitats (Davies *et al.* 2005; Brereton *et al.* 2007). Similar increases in larval food webs of the Marsh fritillary after the introduction of HLS management on sites can be attributed to the increase in suitable host plants (Ellis *et al.* 2012). Potts *et al.* (2009) show that, whilst the addition of a wildflower mix to grass margins increases adult butterfly abundance, there was a negative effect on the abundance of butterfly larvae, suggesting trade-offs between provisioning resources for the different life stages.

2.3.2.7 *Gaps in the evidence of AES management effects on butterflies*

Similar to pollinators, most of the research literature on butterflies has focused on the response of adults; just two studies reviewed have explicitly recorded the abundance of larval stages (Kruess & Tscharntke 2002; Potts *et al.* 2009). The empirical research addresses a smaller range of AES options than the expert species summaries, and more studies are focussed on options that create floral resource patches, rather than management of existing habitats such as grassland or heathland. However, there are some exceptions (e.g. Section 2.3.2.3.5), so the bias towards floral resource studies is not as pronounced as for pollinating insects. As discussed above (Section 2.3.2.6), only a few studies address limiting factors for butterflies in the context of AES.

2.3.3 *Moths*

Fourteen studies present empirical results for the response of moths, as a discrete group rather than in combination with butterflies or pollinating insects, to AES or related interventions. A further 16 Natural England CS factsheets were available for specific moth species.

Nine studies were specific to the UK, eight in England and one in Scotland. Two were conducted in Germany and one each in the Netherlands, Finland and Sweden. Six studies tested grass buffer margins and hedgerow management options on moths, and all but one presented direct tests of English AES options, with another focusing on both grass buffer strips and hedgerows under Scottish AES prescriptions (Figure 10). The same Scottish study also provided AES-specific tests of grassland management, species-rich grassland management and water course and erosion management options (Fuentes-Montemayor *et al.* 2011).
2.3.3.1 Metrics and taxonomic coverage of moth studies

Eleven studies present results for aggregated moth abundance, five for species richness. A common distinction across studies is between macro-moths and micro-moths, with many only presenting results for macro-moth species. Five studies present results for individual species, including five section 41 species (see Figure 11), with one, the Pale shining moth (*Polia bombycina*) being the focus of a research paper aimed at addressing the utility of non-tailored AES interventions for rare species (Merckx *et al.* 2010b).

2.3.3.2 Spatial and temporal coverage of moth species

All studies were spatially replicated, although a large proportion were replicated over a number of years: two studies used data collected over two years (De Snoo *et al.* 1998; Hahn *et al.* 2015), two over three years (Merckx *et al.* 2010b; Staley *et al.* 2016) and two studies over four years (Merckx *et al.* 2012b; Korpela *et al.* 2013). However, only Taylor and Morecroft (2009) present temporal data, assessing changes in moth species richness and abundance (at various taxonomic resolutions) over a twelve period after the start of organic management, although this study only considered a single farm.

Figure 10 Summary of the number of studies which tested broad AES option grouping for moths and the proportion in these groups covered by different study types. AES CF = direct test of AES using a counterfactual control; AES NCF = direct test of AES using a non-counterfactual control; NAES CF =
test of AES relevant habitat using a counterfactual control; NAES NCF = test of an AES relevant habitat without a counterfactual control.

2.3.3.3 Evidence for response of moths to AES management from the research literature

Given the small literature available for review on the response of moths to AES management, any conclusions derived are tentative, and focus on hedgerow management and grass buffer strips as studies assessing these options dominate the evidence base (Figure 11).

2.3.3.3.1 Hedgerow management

Figure 11 Summary of the response of measures of total abundance of moths to AES option groups, under the different study type classifications. AES CF = direct test of AES using a counterfactual control; AES NCF = direct test of AES using a non-counterfactual control; NAES CF = test of AES relevant habitat using a counterfactual control; NAES NCF = test of an AES relevant habitat without a counterfactual control. A non-significant effect was recorded when a result was not significant at \( p \leq 0.05 \), but was significant at \( p \leq 0.1 \). Green bars show the number of recommendations within specific option groups across all available species based summaries of expert opinion.

Whether hedgerow management has an effect on moths is dependent on the AES option under consideration, interactions with management of other habitat features and moth functional grouping. The presence of hedgerow trees in fields increased the abundance of macro-moths that
feed on grass or herbs, but had no effect on the abundance of those species that feed on shrubs or trees (Merckx et al. 2010a), possibly because grass- or herb-feeding species have lower mobility and thus are more dependent on the shelter provided by isolated trees (Merckx et al. 2010a). Merckx et al. (2012b) found positive effect of hedgerow tree presence on total moth abundance, but not species richness. Further, Merckx et al. (2009b) found that positive effects from hedgerow trees on the abundance of large moth species were not significant on their own, but were when applied in combination with 6m grass margins (ELS option EE3), and in landscapes where AES option uptake was actively encouraged. Staley et al. (2016) showed no difference in abundance or species richness of moths on non-AES hedgerows and those cut once every two years in the autumn (ELS option EB2), whereas hedges cut every three years in winter (which fall under ELS option EB3) had significantly higher abundance.

2.3.3.3.2 Grass strips and other field margin treatments

![Figure 12](image.png)

Figure 12 Summary of the response of moth species richness to AES option groups, under the different study type classifications. AES CF = direct test of AES using a counterfactual control; AES NCF = direct test of AES using a non-counterfactual control; NAES CF = test of AES relevant habitat using a counterfactual control; NAES NCF = test of an AES relevant habitat without a counterfactual control. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$.

Results indicate positive effects of grass field margin options on moths, although these effects vary with regard to the specific option and group of species under consideration. For example,
Merckx et al. (2012b) found that 6m grass field margins (ELS option EE3) had a significant, positive effect on the number of macro-moth species compared to 1m field margins, but there was no significant difference in macro-moth abundance. Merckx and Macdonald (2015) suggest, based on the results of mark-release-recapture experiments, this was due to 6m field margins supporting more sedentary species. A recent study of the response of macro-moths to AES margin interventions (most commonly 6m grass buffer strips and margins with nectar flower mixes) in four landscapes in Hampshire, found that AES margins supported significantly higher abundances of macro-moths than control margins (1.23 times greater) and the field centre (2.94 times greater).

Hahn et al. (2015) found that conservation headlands without fertiliser input or pesticide input supported a higher number abundance of moths in the families Geometridae and Noctuidae than those with inputs, although the removal of herbicide inputs had no effect. Alanen et al. (2011) found that there was a significant increase in the abundance and species richness of diurnal moth species associated with wildflower sown set-aside plots, with success particularly pronounced for moths that fed on leguminous plants in their larval stage.

2.3.3.3.3 Grassland management option groups

The few studies available indicate that moth species richness, and in some cases abundance, may be promoted by reduced-intensity grassland management. Littlewood (2008) found that plots with low-intensity sheep grazing on a Scottish upland estate supported significantly higher moth abundance and species richness than plots with high-intensity grazing, although ungrazed meadows performed significantly better than either. Fuentes-Montemayor et al. (2011) showed that species-rich grassland on land under AES management supported a greater abundance and species richness of micro-moths, and greater species richness of macro-moths.

2.3.3.3.4 Response of macro- vs micro-moths and individual moth species to AES

Several of the studies above just assess macro-moth species responses, but there is limited evidence from one study that micro-moths may be more sensitive to AES management, at least at small spatial scales, than macro-moths (Heard et al. 2012). The results for individual moth species come from just four studies, but all suggest that AES management can have significant effects on the local abundance of individual species. Merckx et al. (2009a) compared the abundance of nine moth species and found significant positive effects of 6m grass buffers compared to 2m cross-compliance buffers for the Treble lines (Charanyca trigrammica) and Brown-line bright-eye (Mythimna conigera), but no effect for seven other species, including two section 41 species. Merckx et al. (2010b) found no significant effect of 6m field margins vs cross compliance margins on the abundance of the section 41 Pale shining moth (Polia bombycina), but that hedgerow trees had a significant, positive effect on its abundance. Four of 14 species were found to have significantly lower abundance in fallow fields than in margins in a Finnish study (Kuussaari et al. 2007) and the Lunar underwing (Omphaloscelis lunosa) increased significantly in abundance after an AES and organic farming intervention on a site in Oxfordshire (Taylor & Morecroft 2009).
2.3.3.4 Evidence for the response of moths to AES management from expert species review

Natural England CS factsheets are available for 17 species of moth, 16 of which are section 41 species. Threatened and priority species management are the most abundant option group for management recommendations (Figure 11). For other options, as with butterflies, the diversity of option grouping recommendations covered is much larger than that represented by the empirical evidence, with species-rich grassland, other grassland management actions and woodland management options dominating, linked to habitat specialisms and larval host plant requirements of many species.

2.3.3.5 The influence of landscape on responses of moths to AES management

Four studies considered how landscape variables affect the response of moths, which together suggest that both agricultural practices in the wider landscape and the availability of semi-natural habitats can modify responses to AES, in what may be taxon specific ways. Merckx et al. (2009b) tested the effects of the level of AES uptake within landscapes surrounding hedgerow trees, and found macro-moth abundance was greater for farms in landscapes where AES uptake was encouraged. However, in assessing the effect of grass margins and hedgerow trees, Merckx et al. (2012b) found that the amount of arable land in the landscape surrounding farms (at radii of 1-8km) had no effects on either the abundance or species richness of macro-moths.

A study on the effects of AES interventions on moths in Scotland found that species richness was significantly, positively related to the proportion of semi-natural habitat in a 1km buffer, although it explained a small percentage of the variation (Fuentes-Montemayor et al. 2011). For macro-moths, both species abundance and species richness were positively, but not significantly, related to the proportion of semi-natural habitat in a 250m buffer around a site. The recent work of Alison et al. (2016) shows that the abundance of macro-moths associated calcareous grassland responds positively to AES margin interventions, but also that this increase in abundance is greater when connectivity to calcareous grasslands was higher. This interaction was not found for species which are associated more broadly with grassland or for those not associated with grassland.

2.3.3.6 Gaps in the evidence of AES management effects on moths

There is a paucity of evidence available for moths, with just a few studies dominating the results for this group. As such, all of the evidence gaps present in the literature for pollinating insects and butterflies are applicable (few AES options groups being addressed, no direct information of limiting resources and a focus on adult life stages) but to a more extreme degree and with even less consideration of landscape-scale effects. However, in addressing the application of 6m field margins and hedgerow trees together, the work of Merckx et al. (2009b) is the only study reviewed across all insect groups which has considered how benefits might occur when two different types of AES options are applied together.
Beyond the pollinator, butterfly and moth section 41 species covered above, very little literature was found for other section 41 invertebrate species. Although no scientific papers were found that met the literature review inclusion criteria, three papers present research of some relevance focusing on section 41 invertebrate species which are thought likely to be responsive to AES.

A study from the Czech Republic examining the effects of physiochemical and environmental variables on the presence of the damselfly *Coenagrion ornatum* found that its occurrence was positively correlated with a high diversity of macrophyte vegetation, but were negatively correlated with shading and gravel or concrete substrates (Harabiš *et al.* 2015). Watson and Ormerod (2004) investigated the distribution of three section 41 snail species (*Segmentina nitida, Anisus vorticulus* and *Valvata macrostoma*) in drainage ditches in SE England, in relation to water quality, vegetation and anthropogenic factors. The distribution of *S. nitida* was associated with shallow, calcareous ditches with dense emergent vegetation, *A. vorticulus* with less calcareous ditches and high plant diversity and *V. macrostoma* with floating plants and slightly elevated chloride levels. Eutrophication, in the form of elevated nitrate and nitrite levels was associated with the absence of *S. nitida* and *V. macrostoma* from otherwise suitable ditches. Importantly, the authors suggest that conservation of these species requires quasi-traditional, rotational ditch clearance at the scale of individual marshes, and site protection and catchment fertiliser reduction at a regional scale. In a survey of invertebrate species associated with soft cliffs, Howe (2015) show that the distribution of the section 41 wasp *Odynerus melanocephalus* has increasingly retreated to soft cliffs or shows general regional restriction to soft cliffs in the UK.

Many section 41 invertebrate species have restricted distributions and low abundance, making them poor candidates for a multi-taxon landscape monitoring scheme of AES interventions. As discussed at the workshop (Section 2.2) and in the pollinator and butterfly reviews above, rare or restricted section 41 invertebrates may be better monitored through targeted, bespoke surveys focussed at the sites on which they occur. An alternative to monitoring their abundance might be to analyse changes in distribution through occupancy modelling, although this has drawbacks in the context of monitoring the effects of a specific AES scheme over a 5 year time period, discussed further in the analytical review (Section 5.3.5.2).

### 2.3.5 Summary of invertebrate evidence review

#### 2.3.5.1 Evidence for landscape impacts of AES and landscape scale effects on invertebrates

In the available evidence, the concept of the landscape in the response of invertebrates to AES has been addressed in several ways (as discussed in Section 1.1): (1) whether characteristics of the landscape surrounding farms/options alter the responses of taxa at the local scale to AES, (2) assessing effects of the presence or quantity of AES interventions in the surrounding landscape on short-term, localised response of taxa, (3) assessing effects of existing and previous AES coverage both at a monitoring site and/or in the wider landscape on long term populations responses and (4) assessing whether local scale AES interventions affect the abundance of
organisms in the surrounding landscape not under AES management (potential spill-over effects).

In the context of landscape-scale monitoring of the effectiveness of AES (4) above may be the most relevant, and among invertebrates this has been addressed for pollinators. This approach gives the best indication of how local scale AES interventions are affecting biodiversity, and potentially ecosystem service delivery, in surrounding agricultural land. Importantly, it also indicates whether changes in abundance at the local scale reflect a redistribution of individuals towards resources provided by AES interventions, or changes across wider landscape units including areas without AES interventions. Only recently has literature emerged addressing this issue, and specifically for pollinating insects (Carvell et al. 2015; Jönsson et al. 2015; Scheper et al. 2015), but it is a recognised priority for research.

Point (3) has been addressed by three studies on butterflies in the UK, as this is the only invertebrate group reviewed for which data from long term monitoring of abundance is available at a large number of sites (Brereton 2002, 2005; Oliver 2014). However, these studies demonstrate the potential utility of such data in monitoring how the application of AES in landscapes surrounding sites might influence national temporal population trends, and highlight the potential benefits of developing new methodologies to do this. Point (1) above is addressed most often in the studies reviewed, as many are concerned with understanding how the local level effectiveness of AES interventions will be modified by the surrounding landscape. Whilst, in general, results suggest that qualities of the landscape surrounding sites will influence AES outcomes, it should be noted that there is considerable variation in both the definition of what constitutes the landscape scale and how landscape characteristics are quantified. Several large scale or meta-analysis studies on insect pollinators provide some of the most rigorous evidence for point (1) (Section 2.3.1.3.4) and all suggest that local scale responses to AES will be strongest in ‘simple’ landscapes; e.g. those which have > 1% but < 20% semi-natural habitat coverage, as opposed to ‘cleared’ at < 1% and ‘complex’ at >20% (Scheper et al. 2015). Point (2) has only been addressed for pollinators, and by a couple of studies (Pywell et al. 2006; Scheper et al. 2015), discussed above (Section 2.3.1.5)

In conclusion, there is a general lack of evidence on the relationship between changes in local abundance due to AES interventions and abundance or species richness in the broader landscape (point 4), measurement of which would be a primary goal of landscape monitoring. The lack of evidence suggests that the extent of these effects in practice from future management, such as under CS, cannot currently be predicted.

2.3.5.2 Evidence for local scale effects of AES interventions

For insect pollinators and butterflies there is strong empirical evidence for local-scale responses to AES or AES-like interventions, for some species and in aggregate groups (e.g. all pollinators, all bees, bumblebees, all butterflies). There are biases with regard to the distribution of results among the types of AES interventions; whilst there is some strong evidence for the effects of field margin interventions (particularly those utilising wildflower mixes), floral plots, reduced input and low-intensity grassland management, organic farm management (reduced pesticide, herbicide and fertiliser inputs) and general AES interventions, there is little research evidence
available for other option types, many of which are likely to impact these target groups. There are taxonomic biases amongst pollinating insects with regard to the focus of studies. Bumblebees are assessed far more often than either hoverflies or solitary bees as a discrete group. Almost all studies have focused on testing effects on the local abundance of adult insects, rather than larval / pupal stages, and with regard to exploring which resources are provided by AES interventions, flower resources are by far most commonly quantified. For moths there is a small amount of research evidence available on the impacts of AES interventions and interactions with landscape factors, which provide some comprehensive evidence relating to the impacts of ELS and HLS hedgerow management and 6m grass margins / buffer strips, but less evidence for other AES option groups.

Within each taxon, invertebrate results are dominated by response variables that consider abundance aggregated across species within taxonomic groupings, or species richness. Some evidence is also available for the response of individual species, where studies have looked at responses at various taxonomic resolutions. However, studies with results at species level are generally restricted to bumblebees or butterfly species, and these results are generated by a relatively small number of studies which address a limited range of AES options. Within the species resolution results, 19 section 41 invertebrate species are represented: 11 butterflies, three bumblebees and five moths. Among these 19 species, findings are restricted to a small number of option groups and for many species the number of results is also low. In some instances, species richness was found to respond to AES intervention where abundance of individual species and aggregated across species did not (Section 2.3.1.3.2).

Amongst the species level results reported for pollinating insects, results are most abundant for more common species (e.g. *Bombus lapidarius*, *B. pascuorum* and *B. terrestris lucorum agg*). This is less obvious for butterflies given that many of the species level results were generated from Brereton (2002) and Oliver (2014) based on UKBMS and WCBMS data, which by virtue of their large scale sampling effort can cover many species, but still it is the more widespread species for which the most results and significant effects are reported (e.g. Green-veined white, Meadow brown, Ringlet, Wall brown and Small heath). As with pollinating insects, the more common species also appear to be more responsive to AES (Figure 9), but this may be because common and widespread species are more likely to have adequate sample sizes to detect differences. This is likely to be an important general feature for any monitoring programme.

Across pollinating insects, butterflies and moths, studies have primarily been spatial in their approach (comparing AES treatments against controls and / or each other, using replicated plots over relatively short time periods), rather than addressing longer term temporal trends after the implementation of AES. The exceptions to this are Oliver (2014) and Brereton (2002, 2005), in which changes in temporal trends were found to link to AES intervention for some butterfly species. Studies have for the most part used the abundance of adult individuals at the scale of the plots or field (or presence when investigating species richness) as their response variables, which makes it difficult to determine whether short-term changes in abundance are related to sustained changes in populations. Studies providing longer time series together with measurements of abundance across the landscape are more useful in this regard, to demonstrate that the differences between AES options and controls are sustained over time and to infer population spill-over from areas with AES interventions to the wider landscape.
2.3.6 Birds

Figure 13 A summary of the number of references (within and between studies) that test broad AES options on bird abundance (occurrence). This figure shows where the research emphasis has been placed. Apart from the formal literature search covered in Appendix A, references were also obtained during the workshop and consultation period, with invited experts for the various taxa covered. This exercise uncovered not only peer-reviewed papers and project reports, but also advice notes that were considered in this review as ‘expert opinion’ unless supported by a peer-reviewed citation. The review also draws on work done under a parallel review of the evidence of AES effects on birds for Natural England (Siriwardena & Dadam 2015). The reference list mainly pertains to UK studies as these were deemed to be most relevant to UK AES implementation. However, potentially relevant near-continental studies were included, such as the long-running Dutch meadow-bird work on lowland breeding waders (Kahlert et al. 2007). Search emphasis (a further filter) was also placed on studies looking for population change, and ideally, temporal population change at large spatial scales, given that the emphasis of the scoping exercise was to assess the potential to detect landscape scale effects of AES on species abundances. So, the
the literature review for birds does not fully represent all spatial studies especially at small spatial scales and is therefore not exhaustive in that sense, but it covers all relevant studies for large-scale effects. For the purpose of the scoping exercise, the review concentrates the most relevant, key studies published in the last 20 years.

Figure 14 A summary of detected effects of AES option groups or equivalent non-AES habitats on bird abundance or occupancy. This figure shows in which habitats/options effects have been detected. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$.

The summary output from the literature review, based on 89 bird papers, shows that the range of AES options or equivalent habitats (i.e. not normally specific option codes but habitats such as ‘field margins’ or ‘winter bird crops’ or ‘hedgerow management’) that have been investigated to a greater or lesser extent is large (see also Siriwardena & Dadam 2015). There has been a clear emphasis on winter food supplies (winter bird crops and stubbles etc.), and linear features (field margins, buffer strips and hedgerows) in terms of AES responses by birds (Figure 13). Due to the nature of many studies, in which multiple species have been examined often across more than one option/habitat type, there is considerable repetition within the data sets, with mixtures of
non-significant and significant responses, as would be expected of cross-species ‘screening’. So, with many permutations of species versus option type being exploratory, taking an inclusive approach to species richness and not being based on true experiments, the proportion of non-significant responses to AES or equivalent habitats is high, at over 70% in many cases (Figure 14). This is reflected again in the breakdown by bird species studied (Figure 15), with emphasis on the typical farmland specialists that have been the subject of so much research effort.

Figure 15  Number of referenced significant and non-significant effects of AES, or related habitats, on bird temporal change in abundance (by species or species groups). Some studies tested multiple combinations of bird species versus AES option type, so ‘N’ does not represent the number individual studies. This figure shows the species-level split in detected responses.

The high number of non-significant responses is an important consideration for judging how likely it is that any monitoring system will detect changes in response to AES management. A more constructive way of looking at this might be to ask, ‘which studies or field methods did manage to detect an effect (positive or negative)’ and ‘could these be incorporated onto or inform possible a future AES monitoring scheme?’ The number of effectively experimental comparisons, that include controls, is difficult to assess as often these details are not explicitly identifiable within the methods of studies. Generally, management gradients or differences in habitat types have been compared rather than manipulations of management versus true counterfactuals, limiting the power of the tests and increasing the likelihood of confounding factors influencing the results. True experiments may be logistically impossible at the landscape
scale, so the value of a tight, clear, experimental comparison needs to be judged against that of larger-scale studies that are more representative of real populations of birds.

Table 2

<table>
<thead>
<tr>
<th>Conservation headlands</th>
<th>Ditch management</th>
<th>Fallow plots for ground-nesting birds</th>
<th>Grass buffer strips or margins</th>
<th>Hedgerow management</th>
<th>Low input grassland</th>
<th>Low input grassland &amp; species-rich grassland</th>
<th>Pollinator flower and nectar sources</th>
<th>Wet grassland</th>
<th>Winter bird food sources, incl. stubbles</th>
<th>Winter stubble and summer fallow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullfinch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Corn bunting</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Dunnock</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Grey partridge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>House sparrow</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Lapwing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Linnet</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Meadow pipit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Reed bunting</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Skylark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Song thrush</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Tree sparrow</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Turtle dove</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Yellow wagtail</td>
<td>1</td>
<td></td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Yellowhammer</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>--------------</td>
<td>-----------------------------------------</td>
<td>----------------------------------</td>
</tr>
</tbody>
</table>

Table 2 A matrix of the number of referenced significant tests of AES options on bird species temporal abundance (with known counterfactual). This table informs about potential for detecting temporal change in bird populations in relation to habitat/options.

Table 2 presents a summary of the number of temporal studies that have detected a significant response (whether positive or negative) in population change over time, in response to AES or equivalent habitat management, and where a counterfactual or control is evident. This quickly reduces the number of viable studies to just a few and reduces the species list too. Nevertheless, under the right design and structure, there is evidence of large-scale effects on population growth rates in birds in response to AES management (e.g. Baker et al. 2012), analyses of type (3), as defined in Section 1.1. In this regard, birds are suited to temporal sampling at large scales in a consistent, repeatable way and at a scale that represents national population change, as has been demonstrated through existing monitoring under the BTO/JNCC/RSPB Breeding Bird Survey (BBS). At the same time they are amenable to more intensive surveying that may be required to link abundance change to AES habitat management. Going forward, the BBS data can provide important contextual information on population abundance and trends, to underpin more targeted AES effect monitoring at local scales such as immediately around in-scheme farms. Further analyses of BBS data as conducted by Baker et al. (2012) will be valuable in the future for measuring regional and national responses to management, but will not detect landscape-scale spill-over from managed into unmanaged areas efficiently, because survey squares are rarely
very close to one another and because the survey method is ill-suited to the detection of small, local effects. Such analyses will also become less viable if the geographical extent of AES coverage is reduced to small hotspot areas.

2.3.6.1 AES options summary for birds

The following is a brief overview of option and AES effects on bird populations, via the literature review process, and greater detail is provided by (Siriwardena & Dadam 2015).

2.3.6.1.1 Stubble management and stubble management followed by a summer fallow

Cereal stubble is well studied and includes studies of habitat selection in winter at the field and farm scales, as well as of responses of breeding populations at the national scale. Cereal stubble is recognized as both a preferred foraging habitat for granivorous birds in winter and a significant correlate of population growth rates via positive effects on over-winter survival. AES stubble options critically preclude herbicide inputs, to enhance weed seed availability, and ensure stubbles are left unploughed until late winter, so studies of the value of stubble per se are not relevant to assessment of AES impact. Long-term, national-scale monitoring data have been analysed to measure the effects of stubble management and other AES options on bird population growth rates in England and Wales. There is strong evidence that stubble management has a positive impact on farmland bird population growth rates (in one or more landscapes and at one or more scales), probably reflecting the provision of winter food resources that address the key limiting factor for abundance.

To date, there has been little specific monitoring of “extended stubble” management as an AES option with summer fallow, being newly available to schemes only after the demise of set-aside in 2007, although the winter effects of the option should be the same as those of standard stubble management. The evidence for positive effects of this specific management is, therefore, limited to the results of studies of rotational set-aside, which show habitat selection by a range of predominantly granivorous species in both winter and the breeding season. A two-year study on 28 farms piloting management under the Arable Stewardship Pilot Scheme in England showed that Lapwing nests in these fields had greater daily survival probabilities than nests found in other arable fields.

2.3.6.1.2 Wild bird seed mix (WBSM)

In general, there is strong evidence of the association between winter densities of target species and WBSM at the farm scale, and some effects of WBSM on population growth rate too at the landscape scale have also been found. However, some significant associations are negative (potentially reflecting influences of disease transmission or increased exposure to predation), and unforeseen negative effects of this option may be becoming more common. The timing of seed delivery, amongst other things, may be an issue limiting positive option benefits, while net negative effects could show a developing ecological trap, which could arise through the attraction of predators or disease organisms. Further research into the ecology of these novel habitats, which must be developing as different taxa adapt to them, is required.
2.3.6.1.3 Field margins

Studies at farm and field scale have shown a general positive relationship between field margins and bird numbers. However, the generally positive habitat association of farmland birds with field margins has been difficult to demonstrate at population level. There is more evidence of positive association between farmland bird numbers at the farm-scale than at the population level, probably because the management does not address factors limiting population growth. This would be an example of the kind of honeypot effect discussed above (Section 1.1).

2.3.6.1.4 Grassland management

Evidence from across Europe suggests that delayed mowing can be very effective in increasing breeding success of ground-nesting species by reducing mechanical nest or chick destruction. BBS analyses of long-term AES effects in England and Wales, mostly involving fertilizer and stocking reductions rather than changes to mowing regimes, have shown mixed associations between grassland management and population change of target species, but with positive relationships for population growth rates of species such as Linnet and Skylark, possibly Corn bunting too and for Lapwing with wet grassland management (Table 2).

2.3.6.1.5 Arable reversion

There have been few studies of the effects of reversion as an AES option on birds (Wakeham-Dawson & Aebischer 1998), and there is no clear evidence for large-scale effects on bird population change. There is evidence, however, that the landscape-scale heterogeneity arising from the interspersion of grassland habitats between arable ones (and vice versa) is positively associated with population trends for a range of bird species.

2.3.6.1.6 Hedgerow management

There is limited evidence of the impact of hedgerow management under AES at the field scale. Results from UK studies showed a mixed relationship between hedgerow management and bird populations. The more sophisticated analyses found different relationships for hedgerow management and possibly changing relationships over time, but with no clear patterns for positive influences across species. However, there was a clear positive relationship between hedgerow management and population growth rate within Tir Gofal in Wales, with five of eleven species tested showing significant associations with hedgerow management (Dunnock, Greenfinch, House sparrow, Song thrush and Linnet), all of which were positive (Siriwardena & Dadam 2015), and showing that population effects can be detected under the correct sampling regime. Of course, many farmland passerines may not be limited by breeding success or nest site availability, in which case a clear population response to hedgerow management is not necessarily expected.

2.3.6.1.7 Ditch management

The lack of clear evidence of positive effects of ditch management across species could reflect a general lack of studies targeting this habitat and because some species may not necessarily nest
near ditches on farms or indeed use this habitat very much during winter when conditions are more likely to be limiting. However, response effects have been detected for at least two target species that most likely nest along ditches from time to time, i.e. Reed bunting and Yellowhammer, and possibly a third, Tree sparrow, although the mechanism for this species would involve feeding rather than nesting habitat.

2.3.6.1.8 Beetle banks

In general, this option will rarely cover more than very small percentages of fields, so its effect on most species is likely to be limited. Grey partridge, with precocial chicks in mobile family groups requiring foliar invertebrates and abundance limited by chick survival (e.g. Aebischer & Ewald 2010) may be the bird species most likely to benefit from this option, but national-level uptake has been too low to allow definitive tests of effects on populations to date.

2.3.6.1.9 Pollen and nectar and wildflower strips

Overall, there is strong evidence of a positive association between wildflower strips and the diversity and abundance of a range of invertebrates, but limited evidence of effects on taxa relevant to bird diets. There has also been little research on direct effects on bird populations; again, effects are likely to be limited by the area coverage of the option.

2.3.6.1.10 Conservation headlands

There is some evidence of a positive effect of headlands on gamebird productivity and breeding numbers (e.g. (Sotherton 1991). Overall, there is evidence at local scales of positive effects of conservation headlands on gamebirds, which have been the primary target group, but not for other bird species. Uptake has limited the potential to test and to detect effects at larger scales.

2.3.6.1.11 Skylark and lapwing plots

Tests conducted for arable and mixed farms or landscapes reveal a positive effect of skylark plots. Overall, the farm-scale evidence suggests that skylark plots have positive effects on breeding success and territory density, although benefits have been shown to be compromised by co-location with grass margins because benefits for breeding success are more than negated by apparent facilitation of nest predation. Landscape-scale evidence suggests that positive effects may have occurred in mixed landscapes but not arable areas, although low uptake has restricted analytical power. There is some evidence at the farm-scale, but not at population level, that lapwing plots have had a positive association with their target species. These were initially compromised by inappropriate placement on farms, because a lack of advice to farmers often led to their positioning too close to woodland or other vertical structures, which the birds tend to avoid. The subsequent delay in the effective implementation of this management means that the lack of a population response to date may reflect a demographic lag and initial lack of power rather than a failure to address key limiting factors.
2.3.6.1.12 Multiple option and whole scheme tests (‘AES’ category – Figures 13 and 14).

Research considering whole AES agreements at the farm level has considered only “narrow-and-deep” AES management, i.e. locally intensive management at the farm scale (as opposed to “broad-and-shallow” management spread more thinly across a landscape), either with a single-species or farmland bird-community focus. Management for Cirl buntings in south-west England has taken an inclusive approach such that the old CS and HLS agreements implemented there have integrated options to provide all the species’ requirements, as far as possible. Surveys of breeding and wintering birds to reveal habitat use and population changes have revealed strong increases attributable to the AES management but the management has focused on the conservation outcome, rather than on testing individual option effects and attribution to individual options is difficult (e.g. Peach et al. 2001). Thus, the recovery of the species could indicate synergies between different options providing different resources or, more simply, could show effects of the limiting one among the broader suite of options.

A further, medium-term study has examined bird population responses to HLS management in an arable farming area and a mixed farming one, via three-yearly surveys of breeding bird populations at the farm scale. Bright et al. (2015) found that Grey partridge, Lapwing, House sparrow, Tree sparrow, Reed bunting and Yellowhammer increased more on HLS farms than on control farms, while nine other species were non-significant. There was little clear evidence of which AES option types had driven the results because most tests with respect to individual options were non-significant (Bright et al. 2015). Walker et al. (in review) have conducted a further, improved analysis of the same data, with wider countryside BBS data as a control stratum, thus avoiding potential problems with unbalanced treatment and control samples. There were again local population increases over six years in response to HLS management, although many responses decreased in size over time. This could show a sensitivity to weather events in the effectiveness of the options, but it is also likely that a ceiling will be reached in farm-scale abundance as densities rise locally and this may have occurred in HLS farms with high-quality management. Once again, the significant effects of HLS identified at the farm level were not detectable in terms of the resources provided by individual types of management (which were grouped by types of resource provided for analysis). It is likely that power at the option level was low, especially for the winter food options, because many birds using these options will have bred elsewhere. Likewise, many birds breeding on the study farms probably used habitats outside the farm boundary for at least some of the winter and then possibly responded to the AES management via settlement to breed, rather than through a demographic effect. Thus, the results of this study could represent synergies between multiple option types that are not detectable in tests of individual options, but they do not provide strong evidence to that effect.

2.3.7 Terrestrial mammals

The potential to monitor AES effects on mammals is limited and complicated by the specificity of the survey approaches used for different species. Many approaches, such as those based on field signs, are likely to be effective for detecting species’ presences, but rather insensitive to (or unreliable for) changes in abundance. Thus, for mammals, the number of AES option habitats that have been addressed is small with emphasis on grassland and linear features (plus set-aside
e.g. Macdonald et al. 2007). Significant spatial responses are reported for terrestrial mammals, particularly small mammals in grassland, fallow or field margin habitats (Figure 16). However, the majority of small mammal studies have used radio tracking to assess habitat use and so are not suitable or designed to understand population responses to habitat. Research has shown that hedgehogs can use and benefit from key AES management options (Hof & Bright 2010). Also, Brown hare, as a large, diurnal species of open field habitats, is well-suited to monitoring with visual surveys such as the BBS (Noble et al. 2012), but analyses of AES effects have yet to be conducted. There are studies too, of Dormouse responses to hedgerow or boundary management that have shown spatial relationships (Bright & MacPherson 2002), although again, there is no clear evidence of effects on population change.

![Figure 16](image)

**Figure 16** Summaries of the number of significant and non-significant effects of AES or equivalent habitat, for terrestrial mammals. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$.

### 2.3.8 Bats

Only a small number of options (buffer strips, grassland and hedgerows) have been examined for association with bats and one study of grey long-eared bats has produced results, suggesting responsiveness to grassland management (Barlow & Briggs 2012), supporting the potential for bats to respond more generally (Figure 17). Most generic bat monitoring is probably insensitive to the kind of population change effects being considered in the current scoping study. Indeed some existing monitoring schemes may suffer from a lack of reliability in the identification of bats. However, we note a new, remote method of monitoring and identifying trends in bats that uses new generation bat detectors (removing the requirement for observer identification skills in the field, though the analysis of data still requires skilled assessments) to identify localised activity metrics. Some legitimate issues regarding the reliability of static detectors are addressed here (Newson pers. comm.). We note that knowledge of Myotid bats has improved such that we
can now confidently identify a large proportion of recordings to species. Whiskered and Brandt's are an exception, but discrimination of these two species is improving and will continue to do so. Also, while a number of bat species have quiet calls and may be under-recorded, this problem is true of all species groups. For example, there is a big difference in the detection probability of different bird species. Detectability certainly needs to be considered and formally accounted for in subsequent analyses and interpretation, in any form of survey. Currently, Poisson-Poisson mixture models look promising for static bat detectors, and have been applied to similar problems to estimate abundance and detection from replicated counts of animals that cannot be individually identified (Stanley & Royle 2005; Guillera-Arroita et al. 2011).

In relation to concerns that the 'new' technique has not been ground-truthed by catching bats, there is published work showing that the activity of bats is strongly associated with the number of individuals caught through mist-netting (e.g. Kunz & Brock 1975). However, it should be noted that mist-netting is very biased in what species it catches, and within species, and particularly so if acoustic lures are used. Arguably, mist-netting is less likely to produce a reliable measure of relative abundance than static detectors.

This method has been trialled in Norfolk, with some success, and has been used to identify differences in relative abundance (inferred from activity) in response to different land-use types (Fig. 17; Newson et al. 2015). The key issue for detecting relative abundance responses to AES at the landscape scale is then to have sufficient replication of the detectors to detect the effects of more subtle differences in habitats resulting from AES management, and this is where some methodological testing and development is perhaps needed most. Then, associations with population change in species of conservation interest will need to be measured.

**Figure 17** Summary of the number of significant and non-significant effects of AES or equivalent habitat for bats. A non-significant effect was recorded when a result was not significant at \( p \leq 0.05 \), but was significant at \( p \leq 0.1 \).
2.3.8 Herpetofauna (reptiles and amphibians)

Measuring population abundance in a meaningful and comparable way is difficult for most species of reptiles and amphibians. Difficulties and variability in detection present sampling issues while gaining an understanding of population dynamics is especially problematic for species that show large fluctuations in population sizes, making interpretation of any observed variation difficult. There have been significant spatial effects found for this group in response to habitat management, particularly for field boundary options (buffer strips, e.g. Salazar et al. 2016), hedgerows and ditch management) and for amphibians, pond management or creation (Figure 18). Grazing control is regarded as probably being negative or at best a ‘blunt instrument’ for reptile conservation (T Gent, pers. comm.), although no significant effects either way were uncovered in the present review. Some conservation emphasis has been directed towards lowland heathland and sand dune management, not least due to their importance for the range restricted Smooth snake and Sand lizard, for which methods for habitat assessments have been developed (T Gent, pers. comm.). Developments in habitat suitability indices and the use of remote methods of detecting habitat quality (remote sensing, drones, etc.) are considered ways forward for larger spatial scale monitoring for more widespread reptiles and amphibians, although piloting and ground-truthing would be required before they could be applied to the detection of effective habitat enhancement under AES management.

Figure 18 The number of significant and non-significant effects of AES or equivalent habitat, for reptiles and amphibians. A non-significant effect was recorded when a result was not significant at \( p \leq 0.05 \), but was significant at \( p \leq 0.1 \).

2.3.9 Early successional plants including arable species

The initial focus for landscape-scale monitoring of plants proposed by Natural England was on rare arable plants, 15 species (defined as being “of principal importance” in section 41 of the Natural Environment and Rural Communities Act 2006) of which form a rare arable plant
assemblage within the new CS AES. Ten of these 15 species have restricted distributions (classified as nationally scarce or nationally rare, JNCC 2016b), so they are unlikely to occur frequently enough to form part of a national AES monitoring scheme, and are better candidates for local, targeted monitoring at or near sites where they are known to occur. A more widely applicable monitoring approach for arable plants uses vegetation data to attribute scores to species and sites, which describe their relative importance for arable plant assemblages (Byfield & Wilson 2005; IAPA scores). Following discussions with workshop participants (Section 2.2), the focus of the plant literature review was expanded to early successional plant species more broadly. It is recognized that this excludes some perennial dominated plant communities managed under AES options where early successional species are not the target (e.g. species-rich grassland BAP habitats managed under HLS and the higher tier of CS, Stevenson et al. 2005; Critchley et al. 2004b); time constraints precluded a full review of AES management across all plant communities.

2.3.10.1 Geographical, habitat and taxonomic coverage of literature on early successional plant responses to AES management

Sixty-seven reference sources were included in the review of early successional plants, 36 of which contained information from empirical studies (23 sources) or expert opinion (13) solely from the UK. The remaining references were empirical studies from mainland Europe, with nine from the Netherlands, seven from Germany, and two that used data from across five or six European countries including the UK.

Forty-four of the references addressed arable plant communities (including field margins along the edges of arable fields), five were conducted in grasslands, eight along ditch banks, and the remainder were in set-aside, the base of hedgerows, or covered more than one habitat. Seven response variables relating to early successional plants were assessed in the references reviewed. Species richness was the most frequently reported measure (61% of results); abundance was assessed in 24% of cases, percentage cover in 9% and diversity in 4% of reported results, while presence, Ellenberg N score (Hill et al. 2004) and grass / forb ratio were each used in just one study. All empirical studies used survey methods to collect data at species-level resolution for the entire focal plant community or, in three cases, for a sub-section of the focal community (defined lists of 25 – 86 species of specific interest due to their conservation status or role as indicators of a particular habitat type).

2.3.10.2 Evidence for response of early successional plants to AES management from literature review

2.3.10.2.1 Arable plant species

Arable plant conservation under AES management focusses on providing arable habitats that are disturbed through cultivation, with reduced inputs of fertilizer and herbicides and a reduction in competition from crop species. The most frequently tested AES option group was conservation headlands, management of which commonly includes reductions in fertilizer and herbicide application, and is targeted at arable plant conservation as well as at other taxa. Thirty-five of the 52 empirical sources of evidence reviewed found a positive effect of conservation headland
management on arable plants and 12 found no effect (Figure 19). Five instances of negative effects were identified; four were from one study in which the biomass of four arable species increased if fertilizer was added to a headland. The biomass of ten other arable species (including four rare species) was decreased with fertilizer addition in the same study, and arable plant species richness was greater in the absence of fertilizer (Kleijn & van der Voort 1997). The other negative effect compared arable species on conservation headlands with a naturally regenerating margin that was dominated by early successional species, rather than with a non-AES counterfactual control.

Within arable habitats, empirical evidence also showed a positive effect of uncropped cultivated margins and plots options, fallow plots, undersown spring cereals and winter stubble on arable plant species or communities, while expert reviews of section 41 arable species highlighted a role for uncropped cultivated plots, low input cereal and winter stubbles. There was mixed evidence for the role of grass buffer strips in arable habitats. For example, (Marshall 2009) found that total plant species richness was greater along crop edges that were adjacent to grass margins compared to counterfactual sites without grass margins, but the species richness of annual species was lower. The other option group for which a large number of positive effects on arable plants were found is organic farming. Organic management of arable habitats includes reduced inputs of fertilizer and herbicides, and so has similarities with management of conservation headlands within AES, although it is applied at a larger spatial scale.

![Figure 19](image)

**Figure 19** Number of studies and information sources which assessed AES options, summarized in broad option groups, for arable plants. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$. 

45
2.3.10.2.2 Rare arable plant species

Over half of the 71 pieces of evidence for AES option groups on rare arable plant species were from expert species reviews (47) rather than empirical studies (24). This differs to the evidence base for the broader arable weed community which was dominated by empirical studies, and probably reflects the difficulty of studying rare species in replicated surveys or experiments, in comparison to more widespread species or response variables that measure the total arable plant community at a site. Evidence for positive effects of AES management on rare arable plants comes from a subset of the AES option groups covered for arable plants more generally, with most evidence for conservation headlands, uncropped cultivated margins and plots and organic farming. Expert species review covers a wider range of AES options, with a role also suggested for low input cereals, winter stubble and options targeted at threatened and priority species.

Figure 20 Number of studies and information sources which assessed AES options, summarized in broad option groups, for rare arable plant species of high conservation priority. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$.

2.3.10.2.3 Ditch bank species

The majority of the eight studies that assessed the response of ditch bank vegetation to AES management found a positive effect of ditch management (Table 3). The two negative results were from studies outside the UK: grass / forb ratio increased over time under AES ditch bank management in the Netherlands (Blomqvist et al. 2009), and plant species richness reduced...
following fencing of a ditch bank in Ireland (Feehan et al. 2002) (fencing is not a requirement under ditch management options in Environmental Stewardship or CS). Management of adjacent vegetation under AES (conservation headlands, grass or flower rich field margins) had an effect on ditch bank species in some circumstances. For example, plant species richness was increased on ditch banks adjacent to conservation headlands that had not been sprayed with herbicides or pesticides for one crop (winter wheat), but there was no effect of spraying on adjacent ditch vegetation in two other crops (De Snoo 1999).

<table>
<thead>
<tr>
<th>Sig -ve effect</th>
<th>Ditch management</th>
<th>Conservation headlands</th>
<th>Flower rich margins and plots</th>
<th>Grass buffer strips or margins</th>
</tr>
</thead>
<tbody>
<tr>
<td>NS -ve effect</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No effect</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sig +ve effect</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 3** Number of studies and information sources which assessed AES options, summarized in broad option groups, for ditch bank plant species. A non-significant effect was recorded when a result was not significant at $p \leq 0.05$, but was significant at $p \leq 0.1$.

**2.3.10.2.4 Grassland species**

Grassland AES management assessed included creation of new grasslands through sowing of commercial seed mixtures, reduced mowing of hay meadows and reduced fertilizer and pesticide applications. The majority of the evidence reviewed showed positive effects for grassland management (Table 4). Low input grassland management to target Cirl buntings had a positive effect on overall grassland plant species richness, but no effect on parts of the plant community such as forbs (MacDonald et al. 2012). As discussed above (Section 2.3.10), the responses of grassland plant communities to AES management are not comprehensively covered by this review due to the focus on early successional plant species.

<table>
<thead>
<tr>
<th></th>
<th>Grassland management</th>
<th>Low input grassland</th>
<th>Threatened and priority species and habitats</th>
<th>Wet grassland management</th>
<th>AES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sig +ve effect</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Expert species summary</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**Table 4** Number of studies and information sources which assessed AES options, summarized in broad option groups, for early successional plant species in grassland habitats.

**2.3.10.3 Spatial and temporal scale of assessments of AES management effects on early successional plants**

None of the empirical references reviewed focused on assessing plant responses to AES management at the landscape scale, as the main focus of all studies was on local scale
management. Seven of the 54 empirical studies tested whether landscape factors interact with local management to affect early successional plant species at small spatial scales (within field or field boundaries), which relates to point (1) discussed in Section 1.1 above. Most of these landscape factors summarized surrounding land use, e.g. proportion of semi-natural habitat or arable land in the surrounding 2-5 km (Critchley et al. 2004a; José-Maria et al. 2010). Three of the seven studies found no effect of landscape factors on plant community parameters, which were driven instead by localized management (Marshall et al. 2006; Clough et al. 2007; Seifert et al. 2015); two found surrounding land use interacted with local management, often through species pools (Walker et al. 2007; Batary et al. 2011); and two studies showed landscape factors interacted with local management to affect some plant response variables, but not others (Rundlöf et al. 2010; Concepción et al. 2012).

All but four of the empirical studies compared plant responses to AES management with a spatially discrete counterfactual or alternative management treatment assessed at the same time on replicated study units. A minority of these empirical studies presented data from multiple years which were analysed as repeated measures spatial comparisons of AES management and counterfactual / alternative management, rather than as change in response variable over time. Five studies measured change in plant species richness or abundance over time, two of which did not have a comparison with a non-AES counterfactual. Van Dijk et al. (2013) found that indicator species richness was greater on ditch banks entering AES management than on non-AES ditch banks prior to the start of management, but that there was no difference in the change in species richness over time between the two types of ditch.

2.3.10.4 Plant suitability for monitoring to detect AES effects

Vascular plant species are well suited in general for monitoring of AES management effects, as the majority can be identified to species level in the field, and existing sampling approaches are well developed across a range of habitat types (e.g. Common Standards Monitoring approaches for lowland grassland sites with conservation designations managed under HLS or higher tier CS options JNCC 2016a; established survey protocols for arable habitats Walker et al. 2006). The evidence reviewed above suggests that plant communities predominantly respond to localized habitat management, so plants may not be the primary target for landscape-scale monitoring of AES effects, although landscape factors may interact with local management in some cases. Nevertheless, AES management is focused on habitat creation, improvement or maintenance and plant communities are central to ecosystem characteristics and habitat quality for other taxa, so plant monitoring also has a key role in the assessment of AES implementation success within any AES monitoring scheme.

2.4 Habitat and landscape variables that have been previously shown to affect responses of different taxa to AES

Habitat and landscape variables that have been measured or quantified in conjunction with abundance or species richness in the evidence base reviewed here (Section 2.3), are summarized below. Habitat variables (Table 5a) are defined here as those measured mainly during field surveys (or for a minority of habitat variables, extracted from farm management records), while
landscape variables (Table 5b) were calculated using existing spatially referenced datasets (e.g. Land Cover Map, Atlas data, AES uptake data). Such variables are likely to be important in analyses of AES effects either because management action is expected to be dependent upon the habitat context (e.g. to be effective only in arable systems where spring cropping is rare) or effects could be confounded by background conditions (e.g. the management is associated with a particular crop or landscape, so is expected to be more common where that crop or landscape is more common). There strong evidence that the effectiveness of AES-like management is influenced by landscape factors such as the intensity of farming and the density of semi-natural habitat elements present (Tscharntke et al. 2005, Concepción et al. 2008, Section 2.3.5). Specific monitoring and evaluation questions might therefore consider further variables to, or combinations/transformations of, those listed below, based on previous research experience or ecological theory.
<table>
<thead>
<tr>
<th>Habitat variable</th>
<th>Taxa *</th>
</tr>
</thead>
<tbody>
<tr>
<td>Species and abundance of flower visited by insect during survey</td>
<td>PI, Mo</td>
</tr>
<tr>
<td>Floral abundance or density across all flowering plants (e.g. index of floral units)</td>
<td>PI, Bu, Mo, P</td>
</tr>
<tr>
<td>Floral abundance or density per plant species (index of floral units)</td>
<td>PI, Bu,</td>
</tr>
<tr>
<td>Abundance of nectar-producing flowering plants</td>
<td>Bu, Mo</td>
</tr>
<tr>
<td>Abundance of insect-pollinated flowering plants</td>
<td>Mo</td>
</tr>
<tr>
<td>Diversity of flowering plants</td>
<td>Bu</td>
</tr>
<tr>
<td>Species richness and cover of all vascular plants</td>
<td>PI, Bu, Mo, Bi</td>
</tr>
<tr>
<td>Species richness of forbs and flowering plants</td>
<td>PI, Bu,</td>
</tr>
<tr>
<td>Number, density or presence of larval host plants</td>
<td>Bu, Mo</td>
</tr>
<tr>
<td>Species richness and cover of specific weed species or groups (e.g. broad-leaved weeds, 'key' weed species)</td>
<td>Bi, Ma</td>
</tr>
<tr>
<td>Seed availability or density (e.g. seedhead density, index of seed-rich habitats)</td>
<td>Bi</td>
</tr>
<tr>
<td>Vegetation micro-climate (temperature)</td>
<td>PI, Bu, P</td>
</tr>
<tr>
<td>Vegetation height</td>
<td>PI, Bu, Mo, Bi, Ma, P</td>
</tr>
<tr>
<td>Bare ground cover / area</td>
<td>PI, P, Bi, Ma</td>
</tr>
<tr>
<td>Litter cover</td>
<td>PI, Bu, P</td>
</tr>
<tr>
<td>Shading</td>
<td>PI, Bu,</td>
</tr>
<tr>
<td>Shelter</td>
<td>Bu</td>
</tr>
<tr>
<td>Insolation / sunshine hours</td>
<td>Bu, P, H</td>
</tr>
<tr>
<td>Presence, width or type of field margins</td>
<td>Bu, Ma</td>
</tr>
<tr>
<td>Presence, height or gappiness of hedgerows</td>
<td>Bu, Bi, Ma</td>
</tr>
<tr>
<td>Density or age or management of wet habitat features (ponds, drains, pools)</td>
<td>Bi, Ba, H</td>
</tr>
<tr>
<td>Crop type in adjacent or focal fields</td>
<td>PI, Bi, P, Ma</td>
</tr>
<tr>
<td>Soil variables (e.g. pH, P, K, N, C, Mg, Al, Fe, Mn, B)</td>
<td>P</td>
</tr>
<tr>
<td>Precipitation</td>
<td>P</td>
</tr>
<tr>
<td>Invertebrate abundance</td>
<td>Bi</td>
</tr>
<tr>
<td>Distance from nest to foraging site</td>
<td>Bi</td>
</tr>
<tr>
<td>Habitat / block shape</td>
<td>Ma</td>
</tr>
<tr>
<td>Management (e.g. grazing regime)</td>
<td>PI, Bu, H, Ba</td>
</tr>
<tr>
<td>Farming intensity (measured in various ways including nitrogen input, livestock density, pesticide input)</td>
<td>PI, Bu, P</td>
</tr>
</tbody>
</table>

**Table 5a** Habitat variables (usually measured with a field survey or from farm management records) analysed in conjunction with taxa response variables (species richness, abundance, density, cover etc) to AES interventions in the literature reviewed above (section 2.3). Taxa * PI = pollinating invertebrates; Bu = butterflies; Mo = moths; Bi = birds; Ba = Bats; Ma = terrestrial mammals; H = Herpatofauna; P = Plants
<table>
<thead>
<tr>
<th>Landscape variable</th>
<th>Taxa *</th>
</tr>
</thead>
<tbody>
<tr>
<td>AES uptake (individual options, groups of options or whole scheme)</td>
<td>PI, Bu, Bi, Ba</td>
</tr>
<tr>
<td>Distance to nearest AES patch (e.g. floral patch, boundary management)</td>
<td>PI, Bu, P</td>
</tr>
<tr>
<td>Farming intensity (measured in various ways including or habitat heterogeneity)</td>
<td>PI, Bu, P</td>
</tr>
<tr>
<td>Habitat heterogeneity - diversity index of habitat types (e.g. grassland, arable, woodland, scrub) either at farm scale or in defined radius around site</td>
<td>PI, Bu, Mo, P</td>
</tr>
<tr>
<td>Habitat connectivity - e.g. length of boundaries within 500m radius with semi-natural vegetation</td>
<td>PI, Bu, Mo, H</td>
</tr>
<tr>
<td>Presence / absence of specific habitat types (e.g. arable, grassland, woodland, scrub, isolated trees) in defined radius (e.g. 200m - 3km) around site, or adjacent to site</td>
<td>PI, Bu, Bi, Ba</td>
</tr>
<tr>
<td>Amount of specific habitat types (e.g. arable grassland, woodland, scrub) in defined radius (e.g. 200m - 3km) around site, or on site</td>
<td>PI, Bu, Mo, Bi, Ma, P</td>
</tr>
<tr>
<td>Amount of all semi-natural habitats in defined radius (e.g. 200m - 3km) around site</td>
<td>PI, Bu, Mo</td>
</tr>
<tr>
<td>Number / diversity of patches of different habitat types (e.g. arable, grassland, woodland, scrub) in defined radius around site</td>
<td>PI, Bu, P</td>
</tr>
<tr>
<td>Fragmentation (perimeter to area ratio of habitat types)</td>
<td>PI</td>
</tr>
<tr>
<td>Distance to nearest patch of specific habitat type (e.g. nearest woodland)</td>
<td>Bu</td>
</tr>
<tr>
<td>Length / area of field boundaries or fields</td>
<td>Bu, P</td>
</tr>
<tr>
<td>Field size (mean or focal field)</td>
<td>Bu, Bi, P</td>
</tr>
<tr>
<td>Altitude</td>
<td>P</td>
</tr>
<tr>
<td>Aspect</td>
<td>P</td>
</tr>
<tr>
<td>Slope</td>
<td>P</td>
</tr>
<tr>
<td>Species pool</td>
<td>P</td>
</tr>
</tbody>
</table>

**Table 5b** Landscape variables (usually measured using spatially references datasets, e.g. land cover map, AES option uptake, atlas data) analysed in conjunction with taxa response variables (species richness, abundance, density, cover etc) to AES interventions in the literature reviewed above (section 2.3). Taxa * PI = pollinating invertebrates; Bu = butterflies; Mo = moths; Bi = birds; Ba = Bats; Ma = terrestrial mammals; H = Herpatofauna; P = Plants
2.5 Existing monitoring schemes and datasets that could contribute to landscape monitoring of AES

2.5.1 Butterflies

Two well established schemes monitor trends in the status and abundance of butterflies in the UK, the Butterfly Monitoring Scheme (UKBMS; http://www.ukbms.org/) and the linked Wider Countryside Butterfly Survey (WCBS; http://www.ukbms.org/wcbs.aspx) which was launched more recently to survey wider countryside butterfly species (Roy et al. 2007). The UKBMS has been in existence since 1976 and the WCBS since 2009. Both schemes utilise volunteer recorders to collect data, but the UKBMS sites are selected by volunteers and biased towards coverage of semi-natural habitat, while selection of 1km WCBS squares is based on a randomly stratified approach that ensures representative coverage of habitats across the wider countryside, although at a lower intensity (Brereton et al. 2010).

The methodology for both butterfly survey schemes is based on a version of the Pollard transect method (Pollard & Yates 1994), which records the number of each species of butterfly seen within an imaginary 5m² recording box moving along a 1km transect. Transects are located along linear features, along which the abundance and species richness of butterflies is greater compared to open habitats (Brereton et al. 2010). Two 1km long transects are surveyed in each WCBS 1km square, while one 1km transect is surveyed at each UKBMS site, both within restricted times of day when butterfly activity is greatest. Transects are walked weekly between April and September for UKBMS sites, whereas WCBS squares are visited two – four times a year, with a minimum of two visits. The minimum of two visits in July and August undersamples some wider countryside butterfly species, for example Orange-tip, Brown argus and Wall brown (Brereton et al. 2010), so four visits are encouraged where possible. In 2015, 1424 sites were sampled as part of the UKBMS and 802 as part of the WCBS (Marc Botham, pers. comm.). Approximately half of WCBS 1km squares are the same as BBS recording squares, and covered by the same volunteer recorder.

Both schemes have been used by studies reviewed here; the UKBMS by Brereton (2005), Davies et al. (2005), Ellis et al. (2012) and Oliver (2014), and the more recently established WCBS by Oliver (2014). The work of Brereton (2005) and Oliver (2014) were specifically concerned with using these datasets for the purpose of detecting responses to AES. Considering AES interventions effects on population change, the analyses of Oliver (2014) was restricted to using the UKBMS dataset as this was the only dataset to provide adequate temporal coverage, but similar approaches may be adopted using the WCBS as more data are accrued.

2.5.2 Pollinating insects

The Bumblebee Conservation Trust (BCT) BeeWalk survey is the first recording scheme in the UK to collect standardised data on the abundance of bumblebee species (https://bumblebeeconservation.org/get-involved/surveys/beewalk/), and has been running since
2011. The scheme is similar to the UKBMS, with volunteers selecting sites and choosing their transect routes. The methodology is also similar with bumblebees recorded within a 4m\(^2\) moving recording box, along a 1km transect. The majority of bumblebees are recorded to species, and where possible the caste (queen, worker, male) is also recorded, but other bee species are not routinely recorded. As of 2015, 281 sites were included in the BeeWalk network.

The recent National Pollinator and Pollination Monitoring Framework (NPPMF) study provided a number of scenarios under which a UK pollinator and pollination monitoring scheme might be implemented, to collect data at the national scale on trends in the abundance and diversity of bees and hoverflies (Carvell \textit{et al.} 2016). Plans for implementation are still in development pending funding, but are likely to include a focus on standardised, systematic sampling alongside ongoing recording and further developing occupancy modelling approaches (Claire Carvell, pers. comm.).

2.5.3 Birds

The principal monitoring data sets for UK birds arise from the BTO/JNCC/RSPB Breeding Bird Survey (BBS) and the BTO Bird Atlas, the last of which was conducted over the period 2007-11 (Balmer \textit{et al.} 2013). The BBS is an annual monitoring scheme covering more than 2000 1km squares in England. Squares are selected at random, stratified by observer density, and surveys are conducted by volunteers, although periodic “top-up” surveys using professional observers have been conducted, especially in remote areas. There is considerable turnover of squares as surveyors leave and join the scheme, but replacement surveyors for “lost” squares are sought and the trend has been for an increase in overall survey coverage since its inception in 1994.

The field method for BBS involves two surveys each year, one between April and mid-May and one between mid-May and June. On each survey, the surveyor follows two 1km, self-selected transects, between 0600 and 1000, recording all birds seen and heard in 200m sections of each transect. Birds are recorded in distance bands so as to allow the estimation of densities using distance analysis. Habitats are recorded by 200m section, including land cover, cropping, field boundary characteristics and features such as woodland understorey density.

BBS data are analysed annually to inform about national trends in common bird species and they have also been used to investigate the impacts of AES options at the national scale (Davey \textit{et al.} 2010a; Baker \textit{et al.} 2012; Dadam & Siriwardena 2014; Siriwardena & Dadam. 2015). The data are well-suited to this purpose because the survey method is low intensity but has high spatial coverage and replication. However, the method is not appropriate for deriving accurate estimates of local abundance, such as would be required to measure landscape spill-over effects of local management, because square-specific annual counts are subject to too much stochastic variation.

BBS data are well-suited to estimating large-scale average relative abundance patterns, as would be suitable for targeting management for widespread species with respect to population centres or for deriving ‘background’ population variables for modelling, such as with models interpolating abundance across all 1km squares using land cover variables and spatial smoothing (Massimino \textit{et al.} 2015). However, atlas data can also be used for this purpose. Bird Atlas 2007-
11 (Balmer et al. 2013) comprised standardized surveys of eight or more 2×2km tetrads in all 10km squares in the Britain and Ireland, so complete relative abundance data are available as a survey baseline at the 10km square scale for all of England. Although at a larger scale than the BBS data, these records are not limited by habitat models to informing about more common species and do not rely on an imputing process to fill spatial gaps. The Atlas dataset also includes maps of the relative abundance of wintering birds measured in the same way.

2.5.4 Terrestrial mammals

Mammal population trend data are collected during bird BBS visits (see above) and allow mammal trends to be calculated for nine of the UK’s easily detectable and widespread species (records of other species are recorded too). The nine species include: Brown hare, Mountain hare, Rabbit, Grey squirrel, Red, Roe, Fallow and Muntjac deer, and Red fox, and while there may be detectability and identification issues to be addressed, abundance trends have been produced (Harris et al. 2016). Observers count live mammals, record signs and use local knowledge, or submit ‘nil returns’. The data provide trends at a country and English region level, where sample size allows, in addition to a UK scale. The scheme can provide background population trends in the farmland context, particularly for Brown Hare. Surveys and monitoring feed into the National Mammal Atlas Project (NMAP), via the Mammal Society (www.mammal.org.uk) which aims to produce a thorough baseline of mammal distribution data. The NMAP use different data sets and expert opinion but largely considers only presence data, so does not include repeat sampling at a local level, and no temporal trend information will be available. Note that the Mammal Society has also published survey guidelines for “Mini Mammal Monitoring” as advice notes, building on the conclusions of Poulton & Turner (2009). These are not currently actively promoted and any data collected currently only feed into the NMAP.

A further Mammal Society monitoring initiative planned for 2017 involves developing volunteer surveys to measure population density in under-recorded areas and species, although to date only initial plans have been made for this work (L. Kubasiewicz, pers. comm.). This will not involve repeat sampling, so could not inform about responses to AES management, but could help to identify areas for more intensive monitoring, or stratification (in advance or post hoc) to inform scaling up of identified AES effects.

The National Gamebag Census (NGC) records abundance and distribution trends for 19 UK terrestrial mammal species and especially Brown hare. The NGC has been run by the Game and Wildlife Conservation Trust (GWCT) since 1961 and collates gamebag numbers of quarry species and predator species reported from 600 estates across the UK. Analysis by the GWCT, commissioned by JNCC, has shown that mammal trends produced by the NGC are representative of mammal population trends in the wider countryside. The NGC is valuable as it provides relatively long-term trends (most from 1961) for many predatory and game species, and has good coverage in remote areas. NGC data provides important background information on the distribution and changes in populations of some mammals.
Across all mammal species, the Mammal Society will soon publish *The Review of British Mammals*, which will incorporate the latest information available on population status and trends. Completion is expected in October (L. Kubasiewicz, pers. comm.). Also Natural England commissioned the Mammal Society to update Harris *et al.* 1995. *A Review of British mammals*. The update has been funded collaboratively between country agencies. The Mammal Society will also be undertaking an IUCN red listing exercise for the species.

### 2.5.5 Bats

There are various bat monitoring schemes, run by the BCT, through the National Bat Monitoring Programme (NBMP). The scheme covers roosts, waterways and hibernacula, as well as field and woodland surveys. Perhaps of most interest for the present report, regarding AES effects, the field and woodland surveys are volunteer-based and use handheld heterodyne detectors to measure presence and activity in a sub-group of bat species (Noctule, Serotine, Common Pipistrelle and Soprano pipistrelle) at 1 km scale, adopting a triangular transect route around the square. The field survey is targeted at more experienced surveyors who are able to identify these four bat species above, in the field – as validating the identification of bats using heterodyne detectors can be difficult. The woodland survey focuses on Barbastelles (though Barbastelles make use of far broader range of habitats than woodland alone and in Norfolk would not be considered a woodland species, where it is most strongly associated with farmland with small pockets of trees; Stuart Newson pers. comm.). Data from four 'core' NBMP surveys are used to produce population trends: roost count, hibernation survey, field survey and waterway survey. The NBMP provides population trends for 11 of Great Britain's 17 resident bat species derived from data collected up to and including summer 2015. Trends are provided at GB-level and also at country-level where sufficient data are available (for England, Scotland and Wales). The total NBMP network now stands at 6201 sites and in 2015, a total of 2017 sites were monitored. The NBMP has the potential to provide valuable background trends to AES-targeted work at the above spatial scales. Indeed, the data from NBMP are currently helping to inform decisions to maintain connectivity across landscapes between sensitive areas and important roost locations – though principally concentrating on Annex ii species (Lesser and Greater Horseshoe, Bechstien’s and Barbastelle Bats). This work has tended to focus on existing SSSI’s, SAC’s and sites close to known hotspots. The landscape work includes: improving the landscape for the rare Lesser horseshoe bat around key sites in Wales (the Landscape for Lessers project), and specific AES work assessing the performance of the original Tir Gofal scheme in delivering benefits for selected bat species in Wales.

The Norfolk Bat Survey was launched by Dr Stuart Newson at the British Trust for Ornithology in April 2013 (Newson *et al.* 2015), to improve understanding of bats and support their conservation. Since this time, the project has analysed over a million bat recordings from across the county and was recently expanded to include the Southern Scotland Bat Survey funded by Scottish Natural Heritage, from May 2016. The NBS uses static bat detectors and software to improve bat identification and detection; the method is less biased than traditional survey methods, and requires less effort or skill on the part of the observers. The collation of data from an 'array' of detectors deployed by volunteers across the landscape provided data on presence,
relative abundance and habitat relationships. This method can be deployed in woodland and indeed there may be a stronger argument for using static detectors for woodland species than walked transects for the following reasons. A static detector will automatically record over a complete night - this level of sampling effort is more likely to detect species with a low detection probability such as Brown long-eared bat, than a walked transect where surveyors can realistically only sample the first part of the night. With walked transects it is not straightforward to compare activity at any two locations along the same transect of between sites, because activity is dependent on the time relative to sunset. Furthermore, sound recordings made during transect surveys are often of poorer quality due to the noise generated by moving. It may be possible to make use of activity data from transects, but the analyses would be far more complex. There are also greater health and safety risks associated with walking a transect at night. Perhaps with the exception of Bechstein's bat, survey work in woodland only, for assumed 'woodland' species may propagate a bias. Similarly the national survey for Daubentons’ bat focuses on waterways, and the National Nathusius' Pipistrelle survey is carried out in September, resulting in more records for this month only. Meanwhile emerging from static detectors for Barbastelle shows that this perceived woodland species occurs widely on farmland too. Responses to AES, which under Countryside Stewardship includes woodland, will require objective methods (Stuart Newson pers. comm.).

2.5.6 Amphibians and reptiles

The National Amphibian and Reptile Recording Scheme (NARRS) targets widespread amphibian and reptiles, and began in 2007, collecting data from over 400 amphibian surveys and 400 reptile surveys within the first six years (by 2012). The scheme is designed to create a baseline from which any future trends in status and distribution can be measured. There is also Common Standards Monitoring guidance for reptiles and amphibians, targeting known sites but potentially able to provided population change across these sites and especially for range restricted animals that receive relatively high focus of survey effort (e.g. Sand lizard or Smooth snake). Wider countryside survey protocols have been developed and tested independently of NARRS (Sewell et al. 2013), and relationships to farmland were recently investigated (Salazar in prep). Environmental DNA sampling is becoming increasingly popular and available for presence/absence surveys for great crested newts for use in national and local contexts. eDNA analysis may also able to provide a more comprehensive inventory of all species present within a sample, so informing about a wider range of aquatic species; however, the current techniques are unlikely to be suitable for quantifying relative abundance.

2.5.7 Plants

The National Plant Monitoring Scheme (http://www.npms.org.uk/) was launched in 2015. Volunteers record botanical data to species level in up to five plots within 1km squares twice a year (late spring or early summer and late summer). Squares that form part of the scheme are selected using a random stratified process, whereby selection is random but weighted towards semi-natural habitats, defined using the Land Cover Map. Selection is conducted within 100 × 100 km regions to ensure even geographical coverage. Over 830 NPMS 1km squares were allocated across England to volunteer recorders between 2015 and June 2016, with data
subsequently being submitted by volunteers for around 40% of the squares allocated. As the NPMS was launched recently, uptake and the number of recorded squares are likely to increase over time.

NPMS survey plots are square (usually 5 × 5m) or linear (25 × 1m), depending on the habitat sampled (linear plots are used for linear habitats such as arable field margins, road verges and hedgerows). Volunteers are asked to record two linear and three square plots in each 1km square surveyed. Coverage of species within the plant assemblage is at one of three levels: inventory level at which all vascular plant species are recorded (in just over a third of all survey squares recorded); indicator level at which a subset of plant species are recorded from defined lists for each habitat type (around 40% of squares surveyed); wildflower level at which fewer plant species are recorded using lists of species chosen for ease of identification.

Countryside Survey (CoS: http://www.countrysidesurvey.org.uk/) is a comprehensive botanical survey with national coverage of habitats across England, which ran at discrete intervals between 1978 and 2007. Professional botanists surveyed up to 343 1km squares once during each of the four years CoS was run. CoS squares were selected using a random stratification process, but unlike NPMS the weighting is to ensure even coverage of habitat types (defined using Land Cover Map categories) across England. CoS survey plots are a combination of square and linear plots depending on the habitat surveyed, with the main ‘X’ survey plots formed from a nested set of square quadrats ranging from 1 × 1 m to 14.14 × 14.14 m. All vascular plants present are recorded in each type of survey plot, and % cover to the nearest 5% recorded to species level at most of the plot types.

Several botanical monitoring projects have collected data on plant communities under AES management options, which may be relevant to the proposed landscape scale monitoring. For example, under the Higher Level Stewardship (HLS) AES research project, 180 farms under HLS management were surveyed once in 2009 / 2010 and again in 2015 / 2016. The sampling design focused on land parcels under AES management as the survey unit, within which quadrats of varying size were surveyed depending on habitat, the most common being 1 × 1m and 4 × 4m (Mountford et al. 2013). Samples of ELS agreements have also been surveyed in national monitoring projects (Boatman et al. 2013a). These AES monitoring projects and CoS may provide valuable contextual data for the proposed AES landscape scale monitoring, but as they are not currently supported by committed funding they cannot be relied upon as contemporary or future data sources.

2.6 Evidence review conclusions

The interpretations of landscape in the context of the response of taxa to AES fall into four broad categories (Sections 1.1 and 2.3.5):
1) How do qualities of the landscape surrounding sampling sites alter responses to AES at the local scale?
2) How do the presence or quantity of AES interventions in the surrounding landscape impact on short-term, localised response of taxa?
2) How do populations within 1 km monitoring squares respond to the presence or quantity of AES interventions both within the monitoring square and at larger spatial scales?
3) How do local scale AES interventions affect the abundance of organisms in the surrounding landscape (potential spill-over effects)?

The evidence reviewed has highlighted differences between the broad species groups in which each, if any, of these interpretations of landscape effects has been addressed. The first question has been most commonly addressed in the evidence reviewed across a range of taxa, including pollinators, butterflies, moths, birds and plants. Whilst this evidence usefully informs decisions regarding the landscape contexts in which AES effort may be best placed to maximise benefits, it does not demonstrate impacts of AES that extend beyond the immediate land under AES management. For mobile organisms, the challenge is to demonstrate that AES benefits extend beyond the short-term redistribution of organisms in response to increased resources supplied by AES management, not least because such redistributions do not necessarily imply that population-level effects will follow. This can only be addressed by testing whether increases in abundance (i) are sustained over time in terms of population growth (or reduced rates of population decrease), or (ii) extend more widely onto surrounding land not under AES management (spill-over effects), or ideally both. Correlative analyses of population changes of some species with respect to AES uptake at local and larger scales have been conducted for butterflies and birds at the national scale (question 3 above), in both cases through the use of long-term monitoring scheme data. Spill-over effects have recently been demonstrated (both onto adjacent land that is not under AES management on the same farm and across the wider landscape) in relation to short-term abundance in three studies of pollinating invertebrates (question 3), but from the evidence reviewed here, not for other taxa.

Existing monitoring schemes for birds and butterflies can provide estimates of national population trends, and more recent monitoring schemes for plants, bumblebees and potentially other pollinating invertebrates will allow population trends to be assessed across a broader range of taxa in due course. These schemes can provide important contextual data on national population abundance and trends, and could be used in any scaling up of results from more targeted monitoring of AES landscape effects. However, existing monitoring sites are not located to maximise the chance of detecting AES effects, as sites were not selected to maximise the contrast in AES uptake along a gradient from high to zero uptake. In addition, current monitoring scheme data are often collected at too coarse a scale (usually across a whole 1 km square), and with methods that are too prone to stochastic variation or turnover of survey areas, to assess spill-over effects of local AES management into surrounding landscapes effectively. Within any given year, some sites monitored by volunteers are not visited, resulting in gaps in the time-series that are less consequential for analyses of national trends, but may be a problem for estimates of populations at regional or landscape scales. Spill-over effects are always likely to be small and difficult to detect in competition with other influences (including any AES management on farms surrounding a focal farm). Hence, high-quality, precise, local measures of abundance and species richness will be required to allow effects to be detected and these are unlikely to be provided by extensive volunteer monitoring as is found in the BBS and WCBS, or is likely to occur under the NPMS or NPPMF.
Almost all of the broad species groups reviewed showed responses to AES interventions, but the evidence base is strongest for pollinators, butterflies, birds and plants. Despite the smaller evidence base, recent work shows moths and bats are likely to be affected by AES, and due to their mobility, may also demonstrate responses at the landscape scale (Froidevaux et al. 2015; Newson et al. 2015). Monitoring of mammals, reptiles and amphibians in the context of AES has focussed more on presence/absence than abundance, making it harder to demonstrate AES effects on population change for these taxa. No empirical evidence was found for AES effects on other section 41 invertebrate species, the majority of which may be too rare for abundance to be monitored in general monitoring schemes or have restricted distributions, so are probably better suited to targeted, bespoke surveys. Of course, adequate replication and representation will be necessary to generate the statistical power to detect AES effects among the many influences on a species’ occurrence and in the context of landscapes (see Section 5.5). There is a large evidence base that showing plant species and communities are responsive to AES interventions, which is only partially reviewed here, specifically in relation to early successional species. Effects of the surrounding landscape on the response of local plant species richness and cover to AES have been demonstrated in relation to question (1) above, but the only study on plants to assess temporal change in species richness reviewed found no effect of AES. Plants are less mobile than the other groups reviewed above, so spill-over effects may take longer to become apparent, and none were reported in the literature reviewed. Nonetheless, plants are responsive to AES interventions, many of which are targeted at plant communities, and botanical data forms the basis of many of the habitat variables that have been assessed in conjunction with the response of other taxa to AES (Table 5a).
Designing a landscape monitoring scheme to assess AES effects across multiple taxa (work-packages 3 & 4)

3.1 Key hypotheses and approach to testing biodiversity responses to AES at a landscape scale

The focus for landscape scale monitoring of AES interventions on taxa is in testing whether short-term, localised benefits of AES (i) extend into the wider landscape beyond land immediately under AES intervention, and (ii) are sustained over time (Section 2 for detailed evidence review and discussion). Clearly, this does not obviate the need to monitor responses within managed areas, which are a pre-requisite of any larger-scale or longer-term effect. Nevertheless, consideration of the wider landscape is especially pertinent as the spatial layout of AES schemes change towards more concentrated interventions over smaller areas (AES ‘hotspots’), as to some extent with the new CS scheme, in contrast to low levels of AES management across a large proportion of agricultural land (e.g. ELS). As mobile organisms move across a landscape to utilise resources, either within or between generations, the relationship between local and landscape AES uptake and delivery may impact upon their survival, reproductive success and population growth. Monitoring sites positioned across gradients of local and landscape AES uptake, which provide a contrast in local and landscape AES, would allow the relationships between local AES interventions and interventions at a larger landscape scale to be assessed (Figure 21). For example, high levels of landscape AES intervention surrounding areas with low levels of localised AES intervention would provide a test of potential spill-over of AES benefits, through analyses of the difference in populations between focal and surrounding areas (Figure 22).

**Figure 21** Contrasting gradients of taxon-relevant AES uptake at local and landscape scales. The local gradient is represented by shading from white (low) to brown (high) in the focal 1km squares, the landscape gradient by pale (low) to dark blue (high) in the surrounding landscape sampling units.
Figure 22 Gradient of taxon-relevant AES uptake at landscape unit scale, with zero or low AES in the focal local squares, which could be used to test whether taxa build up on areas under AES management to spill over onto adjacent habitat (in this case the focal square) not under AES management. The potential for or rate of spill-over may also depend on whether taxa have reached carrying capacity in adjacent areas.

One of the challenges in assessing landscape effects of AES is defining spatial scales at which monitoring can be conducted across a range of mobile animal taxa, from invertebrates to birds. In previous tests of AES efficacy, ‘local’ has frequently been interpreted either as land directly under AES option, or whole farms under AES agreement, and ‘landscape’ as areas around a local site ranging from 1km – 10km radii (Sections 2.3 and 2.6). Studies utilising existing monitoring schemes to test for AES effects on temporal change have analysed species data at the 1km square level in relation to AES uptake in the surrounding 0 – 25 km radius landscape (Baker et al. 2012b; Oliver 2014b). For the purpose of constructing contrasting local and landscape gradients in AES intervention, the local scale is defined here as a 1km square, and landscape scale as the surrounding eight 1 km squares around a local 1km focal square, i.e. a 3 × 3 km landscape unit.

While mobile organisms will move outside 3km landscape squares, especially some bird species, the majority of foraging journeys for any given population are likely to be within 3km (Knight et al. 2005; Siriwardena et al. 2006; Siriwardena 2010; Carvell et al. 2012). The local 1km square scale allows comparable data from recording schemes to be used in conjunction with any new data collected (Section 2.5). A local 1km square may contain comparable land both under and outside AES options (depending on its position along a local AES intervention gradient), allowing the testing of within-square AES option vs. counterfactual effects where appropriate for less mobile taxa such as plants, or considering individual life history stages or elements of species ecology (e.g. feeding or nesting) for more mobile ones.

Other factors such as non-AES resource availability (e.g. oilseed rape crops for pollinators (Holzschuh et al. 2016)) and background landscape variables (e.g. proportion of semi-natural habitat, connectivity) may also interact with AES management to affect whether benefits can be demonstrated for taxa (section 2.6), along with AES uptake at larger spatial scales than 3km squares (e.g. Pywell et al. 2006; Baker et al. 2012b). Many of these factors could be assessed as covariates in tests of species response to AES, using existing spatially-referenced datasets such as the Land Cover Map (see analytical review Section 5.1), but this may reduce the analytical power of the AES test itself. Therefore, to maximise the chance of detecting effects of contrasting local and landscape AES interventions on key taxa, ideally the background landscape characteristics of sampling units should be as consistent as possible across landscape sampling units along a gradient, for example with all such sampling units being drawn from a consistent landscape.
We developed evidence-based local and landscape AES gradients, based on AES option group scores for several taxa from the evidence review (Section 3.2 and Appendix C) and uptake of AES options. The potential for National Character Areas (NCAs) to support contrasting gradients of local and landscape AES interventions was assessed (Section 3.3 and Appendix D). NCAs were used as potential regions within which to assess AES intervention along a landscape gradient, because each NCA is a defined, large area that contains consistent landscape characteristics (https://data.gov.uk/dataset/national-character-areas), so should encompass minimal background landscape heterogeneity, and because NCAs have been used as the landscape areas within which priority statements for the new CS AES have been developed (https://www.gov.uk/government/collections/countryside-stewardship-statements-of-priorities).

3.2 Scoring of taxa evidence

Information compiled through the evidence review was used to score AES option groups for pollinators, butterflies and bird functional groups, for all of which strong evidence had been found that AES management altered abundance and/or species richness (Section 2.3). These scores were used to construct evidence-based AES intervention gradients (Section 3.3), as well as providing a summary of the evidence for each taxon. Scores were based on the type of available evidence (e.g. published peer-reviewed studies vs. expert species fact sheets) and the impact of the AES option on the taxon in question (ranging from no effect to significant effects over landscape scales, and from localised abundance to population change). Evidence and impact scores were combined to provide a single score per AES option group for each taxon. Details of the scoring methodology (Section C1) and scores for a wider range of functional groups within the taxa with sufficient evidence (Section C2) are in Appendix C.

A range of butterfly and bird functional groups were initially scored, while for pollinating invertebrates there was only sufficient evidence to score them as one group (Tables C1 and C2). Analysis of the option group scores between taxa showed that in some cases scores were correlated; for example, scores between all butterfly species and both the wider countryside butterfly species and habitat specialist butterfly species (Table C1) were highly correlated (Spearman’s rho = 0.828, P < 0.01 and Spearman’s rho = 0.727, P < 0.001 respectively). The scores were thus simplified across four key groups: all butterflies, pollinating invertebrates, boundary-nesting birds and in-field birds (Table 6), which reduced the number of evidence-based AES gradients to a tractable number.

The weak empirical evidence for AES management effects for bats, other mammals, moths, reptiles and amphibians (Section 2.3) precluded the use of these species groups in the scoring, as any attempt at scoring would have been based on a very limited evidence base, leading to biases if applied in the context of AES uptake of beneficial options. In addition, plants were not scored, as the evidence review had only partially covered the existing AES evidence for plants (Section 2.3.10), and they were not considered the main focus of the design of a landscape monitoring scheme for mobile organisms (Section 2.6). Despite not being included in defining local and landscape AES gradients, several of these species groups have value as potential candidates for landscape AES monitoring (discussed in Sections 2.3 and 2.6), so sampling approaches and resource estimates for each group are provided in Section 4.
<table>
<thead>
<tr>
<th></th>
<th>All butterflies</th>
<th>Pollinating invertebrates</th>
<th>Boundary nesting birds</th>
<th>In field birds</th>
</tr>
</thead>
<tbody>
<tr>
<td>AES</td>
<td>2 2 1</td>
<td>3 2 2</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>Arable reversion</td>
<td>0 0 0</td>
<td>1 2 1</td>
<td>0 0 0</td>
<td>3 2 2</td>
</tr>
<tr>
<td>Beetle bank, invertebrate and bird nesting sites</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>3 2 2</td>
</tr>
<tr>
<td>Butterfly targeted ELS</td>
<td>2 3 3</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Butterfly targeted ESA</td>
<td>2 3 3</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Butterfly targeted OLD CS</td>
<td>2 3 3</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Conservation headlands</td>
<td>3 2 2</td>
<td>1 2 1</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>Ditch management</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>ELS+HLS</td>
<td>0 0 0</td>
<td>3 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>ESA</td>
<td>3 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>3 3 3</td>
</tr>
<tr>
<td>Extensification</td>
<td>3 2 2</td>
<td>3 3 3</td>
<td>0 0 0</td>
<td>2 2 2</td>
</tr>
<tr>
<td>Fallow plots for ground-nesting birds</td>
<td>3 2 2</td>
<td>3 2 2</td>
<td>3 2 2</td>
<td>3 2 2</td>
</tr>
<tr>
<td>Fen</td>
<td>1 2 1</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Field corners</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>Flower rich margins and plots</td>
<td>3 2 2</td>
<td>3 2 2</td>
<td>3 2 2</td>
<td>3 2 2</td>
</tr>
<tr>
<td>Grass buffer strips or margins</td>
<td>3 2 2</td>
<td>3 2 2</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>Grassland erosion management</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Grassland management</td>
<td>3 2 2</td>
<td>3 2 2</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>Hedge and ditch management</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>Hedgerow creation</td>
<td>1 2 1</td>
<td>1 2 1</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Hedgerow management</td>
<td>1 2 1</td>
<td>3 2 2</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>HLS</td>
<td>2 3 3</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Low input cereals</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>3 2 2</td>
<td>3 2 2</td>
</tr>
<tr>
<td>Low input grassland</td>
<td>3 2 2</td>
<td>3 2 2</td>
<td>3 3 3</td>
<td>3 3 3</td>
</tr>
<tr>
<td>Low input grassland &amp; Species-rich grassland</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>2 3 2</td>
<td>2 3 2</td>
</tr>
<tr>
<td>Lowland heathland management</td>
<td>1 2 1</td>
<td>1 2 1</td>
<td>2 3 2</td>
<td>2 3 2</td>
</tr>
<tr>
<td>Lowland raised bog</td>
<td>1 2 1</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td>Moor and heath management</td>
<td>2 2 1</td>
<td>0 0 0</td>
<td>0 0 0</td>
<td>0 0 0</td>
</tr>
<tr>
<td></td>
<td>All butterflies</td>
<td>Pollinating invertebrates</td>
<td>Boundary nesting birds</td>
<td>In field birds</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------</td>
<td>---------------------------</td>
<td>------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>Evidence</td>
<td>Impact</td>
<td>Combined</td>
<td>Evidence</td>
</tr>
<tr>
<td>OLD CS</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Organic</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Pollinator flower and nectar sources</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Rare breeds grazing</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Sand dune creation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Sand dune management</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Scrub management</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Setaside</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Species rich grassland creation</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Species rich grassland management</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Threatened and priority species and habitats</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Uncropped cultivated margins/plots</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Undersown spring cereal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Upland grassland and moorland management</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Upland management</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Wet grassland</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Wet grassland creation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Wet grassland management</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Winter bird food sources</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Winter stubble and summer fallow</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Winter stubble including following silage or fodder crops</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Woodland management and creation</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Wet grassland creation</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Wet grassland management</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Woodland management and creation</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 6 Scores for butterfly, pollinator and bird taxa across all AES option groups for which published or expert opinion evidence was found (Appendix A for evidence review methodology). Scoring for evidence, impact and combined score were attributed as detailed in Appendix C.

3.3 Contrasting local and landscape AES gradients

Current Environmental Stewardship (ELS and HLS) data were used to calculate AES intervention gradients, as data for the new CS AES were not available at the time of this study. Calculating the AES intervention gradients again with CS data is strongly recommended prior to the launch of any biodiversity landscape monitoring scheme, as uptake data are likely to vary in the future, particularly between ELS and the mid-tier of the new CS AES.

Detailed methodology and figures showing the AES gradient calculations are in Appendix D. In summary, contrasting AES gradients at the local (1 km) and landscape (3 km) scales were calculated for each 1 km square within each NCA. Analysis of taxon-specific AES gradients showed strong overlap between gradients for butterflies, pollinating invertebrates, boundary birds and in-field birds in the majority of NCAs (Appendix D Section 2.1), so average AES gradients were calculated. This finding that AES gradients were largely similar across all four taxa means provides support for co-location of sampling units for these taxa across combined AES gradients, potentially saving resources in relation to quantifying habitat variables, assessing AES implementation and managing the field survey logistics (Section 4.4).

There is little indication that AES gradients are strongly correlated with background landscape variables tested within each NCA (Appendix D Section 2.2). This may be partly because high values for the AES gradients are due to varying combinations of the area of land under AES management, the concentration of ES options within a 1km square, and options that score highly due to their weighting by points (Appendix D Section 2.2 for examples). Tests of the response of taxa to contrasting AES gradients should thus be independent of background landscape variables, at least within NCAs. NCAs were defined to encompass areas that share common landscape characteristics, so at a larger regional or national scale, correlations between the AES gradient and background landscape variables may be stronger.

Four AES gradient classes were defined (Appendix D Section 2.3), and the relative position of each 1km focal square and surrounding 3km landscape unit was determined within a matrix of contrasting local and landscape AES gradients. Fewest focal squares provide the contrast between low landscape (3km) and high local (1km) AES gradients (Table 7), notionally the bottom left of Figure 21. Across the landscape AES gradient, cells are overall the most poorly represented in the lowest gradient class, illustrating the prevalence of ES options when averaged across the $3 \times 3$ km scale. This may alter with the new CS scheme, given the focus on clustering agreements.
Table 7 Number of 1 km squares in England with in each combined local (1 km) and landscape (3 km) gradient class. The 3 km gradient is the average of the surrounding eight 1 km squares around a focal square that make up a 3 km sampling unit square, so both gradients are on the same scale. This illustrates the relative abundance of each local and landscape gradient contrast, but does not show the number of potential sampling units due to overlap between 3 km landscape sampling units.

3.31 Potential National Character Areas for proposed landscape monitoring

Landscape sampling units were selected within each NCA as detailed in Appendix D (Section 3), resulting in twenty NCAs that had sampling units in each cell of the contrasting local and landscape AES gradient matrix (Table 9a), three of which (Dark Peak, North Pennines and North Yorkshire Moors and Cleveland Hills NCAs) may be less suitable for landscape monitoring based on AES gradients combined across taxa, as the taxa-specific gradients were less well correlated in these NCAs (Appendix D Table D1). Excluding these three NCAs, seven of the potential seventeen landscape sampling NCAs are dominated by arable land, one by improved grassland and nine are not dominated by a single category; many of the latter have high coverage of both arable and improved grassland (mixed; Table 9b).

Four examples of selected sampling units within NCAs are below (Figure 23); maps of the full range of seventeen selected NCAs are in Appendix D (Figure D6). The Bedfordshire and Cambridge Claylands NCA (Figure 23a) is arable-dominated and the Needwood and South Derbyshire Claylands NCA (Figure 23d) is dominated by improved grassland; the former currently includes two CS facilitation fund agreements. The Cotswolds (Figure 23b) and The Culm (Figure 23d) NCAs are both dominated by a combination of arable and improved grassland land categories, with one CS facilitation fund agreement in the latter.
Figure 23 Examples of NCAs with potential landscape sampling units (3×3km squares with green boundaries) along a matrix of contrasting local and landscape AES gradients. NB Scales differ between the four NCAs.
The twenty selected NCAs are on average larger than other NCAs (ANOVA: $F_{1, 157} = 48.17, P < 0.001$; Table 8B), but do not differ significantly from the other NCAs in terms of cover of major land cover categories (ANOVA: all $P > 0.05$; Table 8A). Species pools are slightly larger for butterflies (GLMs with Poisson error structure: $Z_{157} = 48.17, P < 0.05$) and birds ($Z_{157} = 3.12, P < 0.01$) in the selected NCAs, though the differences in mean number of butterfly and bird species are small (Table 8B). Species pools do not differ between selected and other NCAs for bees ($Z_{157} = 1.83, P > 0.05$).

### Table 8

<table>
<thead>
<tr>
<th></th>
<th>Proportion of NCA area in different land cover categories</th>
<th>NCA size (km$^2$)</th>
<th>Species pool (number of species)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ARABLE</td>
<td>IG</td>
<td>SNG</td>
</tr>
<tr>
<td>Yes</td>
<td>0.41 ± 0.057</td>
<td>0.27 ± 0.033</td>
<td>0.06 ± 0.011</td>
</tr>
<tr>
<td>No</td>
<td>0.36 ± 0.018</td>
<td>0.29 ± 0.012</td>
<td>0.08 ± 0.0074</td>
</tr>
</tbody>
</table>

The potential to test spill-over of taxa from landscape units along the AES gradient into focal 1 km squares with little or on AES intervention (Figure 22) was also explored. Fifty-seven NCAs had at least one sampling unit in each cell of the low 1 km AES gradient categories, and 11 NCAs had at least two sampling units in each cell.
<table>
<thead>
<tr>
<th>Area</th>
<th>1km</th>
<th>3km</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-100</td>
<td>101-1000</td>
</tr>
<tr>
<td>Bedfordshire and Cambridgeshire Claylands</td>
<td>1 2 4 7 5 12 4 8 4 6 5 2 5 12 3 3 83</td>
<td></td>
</tr>
<tr>
<td>Bowland Fringe and Pendle Hill</td>
<td>1 2 2 1 1 4 3 3 2 1 1 1 1 3 2 2 30</td>
<td></td>
</tr>
<tr>
<td>Cornish Killas</td>
<td>2 3 2 8 3 12 3 6 3 2 3 2 4 7 6 4 70</td>
<td></td>
</tr>
<tr>
<td>Cotswolds</td>
<td>2 8 5 5 3 4 16 13 2 7 2 3 1 8 3 4 86</td>
<td></td>
</tr>
<tr>
<td>Dark Peak</td>
<td>1 2 1 1 1 5 3 3 1 3 1 2 3 1 1 1 30</td>
<td></td>
</tr>
<tr>
<td>High Weald</td>
<td>2 2 1 6 1 2 3 6 2 2 1 3 2 5 5 2 45</td>
<td></td>
</tr>
<tr>
<td>Humberhead Levels</td>
<td>2 1 1 3 4 2 2 5 3 2 4 1 1 9 6 5 51</td>
<td></td>
</tr>
<tr>
<td>Low Weald</td>
<td>2 2 3 4 2 6 7 4 8 4 2 4 1 3 2 1 55</td>
<td></td>
</tr>
<tr>
<td>North Pennines</td>
<td>1 6 3 1 1 4 10 8 1 7 2 3 3 7 1 1 59</td>
<td></td>
</tr>
<tr>
<td>North Yorkshire Moors and Cleveland Hills</td>
<td>3 3 2 5 2 3 4 7 2 2 1 5 1 2 1 4 47</td>
<td></td>
</tr>
<tr>
<td>Northern Thames Basin</td>
<td>6 3 5 1 3 4 5 4 5 8 3 3 1 5 1 5 62</td>
<td></td>
</tr>
<tr>
<td>Notts, Derbyshire and Yorkshire Coalfield</td>
<td>2 3 1 3 3 5 3 7 2 5 3 2 2 1 1 46</td>
<td></td>
</tr>
<tr>
<td>Severn and Avon Vales</td>
<td>3 2 6 5 2 7 4 14 3 1 2 5 1 7 4 3 69</td>
<td></td>
</tr>
<tr>
<td>Shropshire, Cheshire and Staffordshire Plain</td>
<td>2 2 4 12 5 3 4 11 12 7 2 6 3 10 14 3 100</td>
<td></td>
</tr>
<tr>
<td>South Norfolk and High Suffolk Claylands</td>
<td>2 2 1 3 2 5 2 8 3 7 1 4 2 16 3 2 63</td>
<td></td>
</tr>
<tr>
<td>South Suffolk and North Essex Clayland</td>
<td>1 1 3 4 6 6 2 18 5 11 4 5 3 13 5 3 90</td>
<td></td>
</tr>
<tr>
<td>The Culm</td>
<td>2 4 2 7 3 9 7 4 7 6 3 3 2 16 3 2 80</td>
<td></td>
</tr>
<tr>
<td>The Fens</td>
<td>3 5 4 1 7 12 8 18 8 8 4 2 10 15 4 2 111</td>
<td></td>
</tr>
<tr>
<td>Trent and Belvoir Vales</td>
<td>1 1 2 3 5 6 3 3 6 2 2 1 1 13 3 3 55</td>
<td></td>
</tr>
<tr>
<td>Wealden Greensand</td>
<td>3 1 3 6 1 3 3 3 1 4 1 3 2 5 4 2 45</td>
<td></td>
</tr>
<tr>
<td>Region</td>
<td>NCA size (number of 1km squares)</td>
<td>Proportion of NCA area in different land cover categories</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>----------------------------------</td>
<td>-----------------------------------------------------------</td>
</tr>
<tr>
<td>Bedfordshire and Cambridgeshire Claylands</td>
<td>2599</td>
<td>0.64 0.18 0.03 0.21 0.00 0.00</td>
</tr>
<tr>
<td>Bowland Fringe and Pendle Hill</td>
<td>738</td>
<td>0.02 0.68 0.13 0.82 0.04 0.05</td>
</tr>
<tr>
<td>Cornish Killas</td>
<td>2222</td>
<td>0.46 0.31 0.04 0.35 0.01 0.01</td>
</tr>
<tr>
<td>Cotswolds</td>
<td>2887</td>
<td>0.49 0.29 0.09 0.38 0.00 0.00</td>
</tr>
<tr>
<td>Dark Peak</td>
<td>867</td>
<td>0.02 0.22 0.17 0.39 0.20 0.44</td>
</tr>
<tr>
<td>High Weald</td>
<td>1752</td>
<td>0.20 0.44 0.04 0.48 0.00 0.00</td>
</tr>
<tr>
<td>Humberhead Levels</td>
<td>1723</td>
<td>0.72 0.09 0.05 0.14 0.00 0.02</td>
</tr>
<tr>
<td>Low Weald</td>
<td>1816</td>
<td>0.31 0.44 0.04 0.47 0.00 0.00</td>
</tr>
<tr>
<td>North Pennines</td>
<td>2137</td>
<td>0.01 0.18 0.20 0.38 0.24 0.55</td>
</tr>
<tr>
<td>North Yorkshire Moors and Cleveland Hills</td>
<td>1661</td>
<td>0.19 0.23 0.12 0.35 0.22 0.24</td>
</tr>
<tr>
<td>Northern Thames Basin</td>
<td>2511</td>
<td>0.35 0.24 0.01 0.25 0.00 0.00</td>
</tr>
<tr>
<td>Notts, Derbyshire and Yorkshire Coalfield</td>
<td>1705</td>
<td>0.33 0.22 0.05 0.27 0.01 0.01</td>
</tr>
<tr>
<td>Severn and Avon Vales</td>
<td>2090</td>
<td>0.52 0.27 0.06 0.34 0.00 0.00</td>
</tr>
<tr>
<td>Shropshire, Cheshire and Staffordshire Plain</td>
<td>3668</td>
<td>0.41 0.41 0.04 0.46 0.00 0.00</td>
</tr>
<tr>
<td>South Norfolk and High Suffolk Claylands</td>
<td>2145</td>
<td>0.73 0.17 0.03 0.20 0.00 0.00</td>
</tr>
<tr>
<td>South Suffolk and North Essex Clayland</td>
<td>3286</td>
<td>0.69 0.16 0.02 0.18 0.00 0.00</td>
</tr>
<tr>
<td>The Culm</td>
<td>2834</td>
<td>0.36 0.45 0.04 0.49 0.00 0.00</td>
</tr>
<tr>
<td>The Fens</td>
<td>3826</td>
<td>0.84 0.07 0.02 0.10 0.00 0.00</td>
</tr>
<tr>
<td>Trent and Belvoir Vales</td>
<td>1777</td>
<td>0.72 0.13 0.04 0.17 0.00 0.00</td>
</tr>
<tr>
<td>Wealden Greensand</td>
<td>1468</td>
<td>0.27 0.29 0.04 0.33 0.02 0.02</td>
</tr>
</tbody>
</table>

| Table 9 Twenty NCAs (a) which contain at least one potential sampling unit within each cell of a matrix of local and landscape AES gradients, based on four gradient classes and (b) land cover categories, number of monitoring scheme squares and species richness for key taxa. IG = improved grassland, SNG = semi-natural grassland, MHB = combined cover of montane, heath and bog land classes. |
3.32 Existing monitoring schemes’ coverage of the local and landscape AES gradients

Existing volunteer monitoring schemes (BBS and WCBS, Section 2.5) have broad spatial coverage in England and have been used successfully in previous evaluations of AES impacts (e.g. Baker et al. 2012, Oliver 2014). There is clear potential to use these data to test for population change in similar ways in the future, i.e. across larger areas than NCAs, or the whole of England, in relation to the AES gradients developed here, or other variables developed from CS uptake or management data. However, as noted in Section 2.5, these schemes are not well-suited to local-scale monitoring because the sampling within them is not intensive enough to derive accurate estimates of local abundance, and because turnover of squares is high in both WCBS and BBS, so data cannot be guaranteed to be collected for any given site across a specific 5-year monitoring period. As a result, application to any new monitoring scheme focused on landscape-scale effects of CS management would best use these data to provide population context, such as the local abundance of target species, which can then assist with scaling-up of NCA-level results to the national scale. For example, given observed responses in an NCA with a particular density of a focal species and a particular landscape type, likely responses in other NCAs can be predicted by scaling according to the analogous density and land cover data there.

Nevertheless, given that part of the remit of this project was to scope existing volunteer monitoring schemes, the coverage of the local and landscape AES gradients by sites in existing schemes (BBS and WCBS, Section 2.5) was explored as detailed in Appendix D (Section D4). The level of coverage and potential bias with respect to AES gradients will be important in determining the reliability of these data for scaling up. No single NCA had WCBS or BBS scheme sites in every cell within the matrix of contrasting local and landscape AES gradients. Eleven NCAs had at least one sampling BBS site in each cell of the low 1 km AES gradient categories (Figure 22), while six NCAs had equivalent coverage by WCBS scheme squares. This confirms that scheme coverage is thus not sufficiently comprehensive enough to contribute to NCA-scale tests of local and landscape AES gradients across the whole matrix.

3.4 Designing a landscape monitoring scheme: conclusions

The potential to construct evidence-based AES gradients at local and landscape scales has been demonstrated based on ES uptake data, which would allow monitoring of the extent of spill-over of mobile organisms into focal 1km squares from surrounding squares, and vice versa, as well as monitoring to assess the interacting effects of AES interventions at local and 3 km spatial scales on changes to abundance or populations within a focal 1km square. Separate gradients based on four mobile taxa (butterflies, pollinators, in-field birds and boundary birds) with sufficient evidence of AES effects were shown to largely coincide within the majority of NCAs, supporting the potential to co-locate sampling units across all four taxa, thereby saving resources. Within individual NCAs there was little evidence that the AES gradients varied consistently with background landscape characteristics, in terms of proportional cover of arable or semi-natural grassland, or landscape diversity, so measurement of AES effects should not be confounded by such variation. This is probably due to NCAs being defined to encompass land with similar characteristics, as intended. AES gradients covering larger spatial scales across multiple NCAs, or nationally, might co-vary more strongly with background landscape characteristics, although this has not been tested specifically.
The potential for seventeen NCAs to support full contrasts in a matrix between local and landscape AES gradients divided into four gradient classes has been demonstrated. The costs and resources needed to support such monitoring (Section 4.4) could be reduced by sampling part of the matrix, depending on which landscape monitoring research questions are considered most pertinent. For example, if the priority of landscape sampling was to assess spill-over onto non-AES land, one row of the matrix could be used (Figure 22). Alternatively, the corners of the matrix would give the strongest tests of contrasts between local and 3 km AES gradients (and thus the relative importance of local and landscape management), although the two middle gradient classes may be more representative of land under AES management across agricultural land more broadly. The highest gradient class may represent squares with high coverage of options that attract high payments, potentially more likely to be in HLS or the higher tier of the new CS scheme. Another option would be to reduce the AES gradients to three classes, which would diminish the ability to discriminate between different levels of AES intervention, but ensure coverage across the full gradients, albeit more coarsely than with four gradient classes.
4 Species monitoring strategy (work-packages 3 & 5)

A detailed species monitoring approach has been developed, based on a landscape study design using $3 \times 3$ km landscape sampling units positioned along a matrix of contrasting AES gradients at the local (1 km square) and landscape ($3 \times 3$ km) scales, as detailed in Section 3. Monitoring protocols are proposed below, having been developed for a range of taxa using information from: research studies relating to AES management reviewed in Section 2, existing monitoring schemes (e.g. Roy et al. 2007), previous AES monitoring projects (e.g. Mountford et al. 2013), and input from taxon experts who attended the workshop (Section 2.2). A framework for rapid assessment of AES option implementation was developed using the combined project team experience from previous AES and Campaign for the Farmed Environment (CFE) research projects (Boatman et al. 2013a; Mountford et al. 2013; Boatman et al. 2015). It is important to note that there have been very different levels of investment and survey development in the past for different taxa, and the extent to which possible approaches have been applied in farmland differs. Hence, knowledge of the sensitivity and power associated with possible approaches to survey birds and butterflies is far greater than that for amphibians, for example, with plants and pollinating insects being intermediates. The same is true about knowledge concerning the effects of subtle variations in survey approaches, such as the timing, length and number of repeat survey visits. As a result, some of the monitoring protocols proposed below are based on wide past experience and can be adopted with a high degree of confidence in their delivery of evidence, whereas others, whilst based on the best available current knowledge, require piloting. The latter are identified in the text and, in some cases, are presented as a discussion of the possible approaches available because considerable methodological development is required before a monitoring can be designed with confidence.

The levels of replication required to monitor various landscape study designs are addressed in Section 4.3, including a detailed power analyses for one potential monitoring scheme design and a discussion of uncertainty around replication estimates. Resource estimates for fieldwork are presented in a modular form for each taxon, and finally monitoring scheme designs are summarised to present the costs and estimates of replication for a range of options in a research question or hypothesis driven framework (Section 4.4).

4.1 Designing a species monitoring approach: protocols to survey individual taxa

4.1.1 Monitoring pollinating insects and butterflies

4.1.1.1 Methods for sampling pollinating insects

Pollinating insects, defined here as all bee and hoverfly species, are a large and diverse group which presents challenges to effective sampling across the entire group. Various active and passive sampling approaches were used in the AES studies reviewed in this study (Section 2.3), the most common method being transect walks (either purely observational or also using a net to catch some insects for identification), with 34 studies adopting a variant of this approach. Pan traps were used in 11 studies, sweep netting in seven, and timed observation plots in three. Other methods (window traps, flight traps and observations of visits to potted phytometer plants) were
used in a single study, while trap nests sample only a subset of the bee community (aerial nesting solitary bees).

Table 10 Comparison of transect and pan trap sampling methods for pollinating insects

<table>
<thead>
<tr>
<th>Group sampled / criteria</th>
<th>Sampling method</th>
<th>Transect</th>
<th>Pan trap</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bumblebees</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Species level</td>
<td></td>
<td>Species level</td>
</tr>
<tr>
<td></td>
<td>for most bumblebee species; some species and castes cannot be distinguished</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comparable abundance and species richness between methods*</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Solitary bees</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Some species distinguished in field depending on transect width; taxonomic resolution dependent on recorder expertise</td>
<td>Species level</td>
<td>Greater abundance and species richness than for transects*</td>
</tr>
<tr>
<td></td>
<td>Abundance comparable with pan traps*</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hoverflies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Some species distinguished in field depending on transect width; dependent on recorder expertise</td>
<td>Species level</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abundance comparable with pan traps*</td>
<td>Greater species richness than for transects*</td>
<td></td>
</tr>
<tr>
<td><strong>Honeybees</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Honey bees under sampled due to social foraging</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Method bias</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sensitive to weather conditions at time of walking; floral resource availability along transect needs to be quantified; smaller species may be under-recorded; habitat structure</td>
<td>Floral resources in the direct locality and in the surrounding area need to be quantified; height in relation to surrounding vegetation; larger species may escape more frequently</td>
<td></td>
</tr>
<tr>
<td><strong>Recorder bias</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recorders expertise catching insects and in discriminating them in the field; taxonomic expertise of the recorder influences which groups are recorded to species; potential bias towards larger and more conspicuous species</td>
<td>Minimal; potential for misplacement.</td>
<td></td>
</tr>
<tr>
<td><strong>Area sampled</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abundance / species richness estimated over a standardised area (typically ranging from 1 m – 5 m width depending on taxon)</td>
<td>Area sampled broadly consistent over time and between traps; area sampled not defined</td>
<td></td>
</tr>
<tr>
<td><strong>Lab work required</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimal; only samples caught in nets; potentially photograph analysis to aid ID even to group level depending on taxon</td>
<td>All samples require sorting and identifying to group and/or species level; sample storage required</td>
<td></td>
</tr>
<tr>
<td><strong>Recorder expertise</strong></td>
<td>Researcher / consultant</td>
<td>Fieldwork – no expertise needed</td>
<td>Labwork – researcher / consultant</td>
</tr>
<tr>
<td><strong>Ease of use</strong></td>
<td>Route easily repeatable, may require adjustment due to access restrictions. Minimal equipment.</td>
<td>Location easily repeatable</td>
<td>Twice as many visits to site needed (set up and collection). Equipment required</td>
</tr>
</tbody>
</table>

* Conclusions regarding relative effectiveness of methods from (Carvell et al. 2016).
Sampling method suitability depends on the taxonomic resolution required, the pollinator groups under consideration and the conditions under which sampling is conducted. Potential biases emerge from interactions between a given method, the behaviour or morphology of the specific taxon, its environment and the person utilising the method (the recorder). Biases can be minimised or offset by appropriate sampling conditions, deploying complementary sampling methods together and recording confounding factors (Westphal et al. 2008), and as long as these biases remain consistent over time, inferences can be made about temporal changes. It is thus important that variation in biases over time is controlled; this is most likely due to variation in environmental conditions or where data collection is highly dependent on the expertise of the recorder. A combined transect and pan trap approach has been used in several studies, and was piloted as part of the recent National Pollinator and Pollination Monitoring Framework project (NPPMF; Carvell et al. 2016).

The quantity of floral resources at or surrounding a sampling location may generate bias in opposite directions for observation (transect) and passive (pan trap) methods. Transects involve observations of insects visiting flowers, and often (especially for butterflies) include counts of insects in flight, therefore pollinator density tends to increase with flower density along a transect. It is possible that pan trap catches could also be increased by higher floral resource levels, but as pan traps function by effectively mimicking a floral resource to attract insects, increasing floral resources in the vicinity of pan traps may act to reduce catch sizes via competition. Evidence is available to support both mechanisms for pan traps, thus it is important that floral resources surrounding pan traps and along transects are quantified at different scales.

The habitat in which transects are located can affect the number of pollinators recorded. Carvell et al. (2016) tested 1m wide pollinator transects placed across the diagonal of a 1km$^2$ sampling unit (regardless of habitat) vs. transects where the location was adjusted to follow linear features and likely pollinator activity, and recorded a greater abundance on transects along linear features for all groups of pollinators (R O’Connor pers. comm.). Given the high likelihood of zero-inflated data for less abundant invertebrate species and groups, which will reduce the power to detect effects of AES interventions, variations in methodology such as the use of a sampling habitat that can increase the number of pollinators sampled may be beneficial.

4.1.1.2 Methods for sampling butterflies

The Pollard transect is a well-established sampling method (Pollard 1977, Section 2.5.1), used by the main butterfly monitoring schemes across the UK (UKBMS and WCBS) and Europe (van Swaay et al. 2012). Adaptations of this approach are often used in research studies to assess effects of habitat or management on butterflies. Given the ubiquity of this method, its success in providing data for other monitoring schemes and in demonstrating effects of AES management on some butterfly species (Brereton et al. 2007; Oliver 2014) it is proposed as the key approach for sampling butterflies. The added benefit of adopting a protocol applied by other schemes is that the data will be broadly comparable, making it possible to integrate datasets, for example for national scaling-up.

The Pollard transect is a multi-species approach, providing records of adult butterflies to species resolution, and with potential for tailored application to particular species or lifecycle stages.
(Pollard & Yates 1994; Roy et al. 2007). For example, a dusk transect has been developed for monitoring the Purple hairstreak (Favonius quercus), a widespread woodland species which is only active later in the day; standardised egg surveys can be carried out on hedgerows for Brown hairstreak (Thecla betulae) and larval web counts used to assess populations of Marsh fritillary (Euphydryas aurinia). For both the Brown hairstreak and Marsh fritillary these methods have been used in research to demonstrate positive impacts of AES management (Ellis et al. 2012; Staley et al. 2014). These species-specific sampling methods are not recommended as a component of the protocol outlined below, which aims to be broad in species coverage, but could be integrated into this broader approach at a later stage.

The Bumblebee Conservation Trust (BBCT) has successfully used an adaptation of the Pollard transect method for its bumblebee monitoring scheme, BeeWalk (http://www.beewalk.org.uk/; see Section 2.5.2 for details). The Glastir Monitoring and Evaluation Programme (GMEP) also uses an adapted Pollard methodology to record the abundance of all bees and hoverflies in addition to butterflies, classifying individuals into broad groupings rather than species (Emmett et al. 2014). Experience from the CEH-led NPMMF project suggests it is not possible to combine sampling of all pollinator groups to species on single, simultaneous Pollard butterfly transects, as their 5m width and focus on insects in flight as well as on flowers precludes accurate recording of smaller flower visitors (Carvell et al. 2016). The transect and approach trialled as part of the NPPMF project used 1m wide transects (Carvell et al. 2016), which are not wide enough to give combined coverage of field edge, field margin and field boundary (e.g. hedgerow). We propose sampling of pollinating insects and butterflies using a combination of transects and pan traps as follows; 5m transects based on the Pollard approach to monitor butterflies and bumblebees, and pan traps for bumblebees, solitary bees and hoverflies. Concurrent sampling of butterflies and bumblebees increases efficiency, through reducing the fieldworker time required and thus reducing costs. Sampling will be repeated at four times per year at monthly intervals from May – August.

4.1.1.3 Sampling pollinators and butterflies within focal 1km squares

Within each focal 1km square, 2 km length of transect will be walked, split into 10 × 200m sections in fixed locations. Six pan trap stations will be used within each 1km square, placed at the start of six of the 10 transect sections. The 10 transect sections will be placed along linear features; some transect sections may be contiguous, depending on the layout of each 1km square.

The 10 transect sections within a 1km square will be located to be broadly representative of the uptake of AES linear feature and whole field options that have been identified as beneficial to pollinators and butterflies (Section 3.2, Appendix C). The presence of beneficial linear or edge and/or whole field options will depend on the underlying agricultural habitat. In arable habitats, the beneficial options identified for pollinators and butterflies are along field edges, in field corners (e.g. flower rich margins, pollen and nectar mixes) or boundaries (e.g. hedgerow management); here termed linear features. Within the square the proportion of linear features falling into AES /non AES management will be estimated and 10 × 200m sections located approximately proportional to this. For example, if 64% of linear feature length in a 1 km square fall under AES management, 1200 m of the 2 km transects (6 of 10 transect sections) and four of the six pan trap stations will be located on these and the remaining four sections and two pan trap
stations along linear features not managed under AES (Figure 24). Note that non-AES counterfactual locations within survey squares will therefore effectively not be covered in squares with very high AES uptake, but that this will reflect conditions on the ground: in such cases, there would be no non-AES locations that would allow such a comparison.

**Figure 24** Pollinator and butterfly survey transect sections for an example focal 1km square where all AES target options are along linear features.

In pastoral, mixed and upland/moorland farming habitats, options identified for pollinators and butterflies include in-field management (e.g. low input grassland), so half the transect sections (total 1km length) and pan trap stations will be located according to AES coverage of boundary options, and half according to in-field options (Figure 25). In both cases the transect section routes can follow linear features (field boundaries, hedgerows, paths), but should include some of the in-field habitat in the 5m sampled, for example when walking through grassland or heathland.

**Figure 25** Pollinator and butterfly survey transect sections for an example focal 1km square where AES target options are a combination of linear feature and in-field options.
4.1.1.4 Sampling pollinators and butterflies in surrounding eight 1km squares

Pollinators and butterflies may operate at smaller spatial scales than most bird species, so it is unlikely to be efficient to attempt to measure spill-over from a focal square far from its boundary. A subsample of the eight 1km squares surrounding the focal square will therefore be sampled, using the same methods as for the focal 1km square. The proximal half of each of the four closest adjacent 1km squares will be sampled (Figure 26) at the same intensity as the focal square, with five 200 m long transect sections and three pan trap stations in each 0.5 × 1 km.

By keeping the sampling density the same in the sub-sampled surrounding squares, this will allow estimates of butterfly and pollinator species richness to be directly comparable between the focal and outer squares. If a lower sampling density were used in surrounding squares, as proposed for some of the other taxa (Sections 4.13 – 4.17), comparisons of species richness at the 1km scale between focal and adjacent squares would need to be corrected for differences in sampling intensities, and thus prone to more uncertainty. Species richness is of less interest than changes in abundance for taxa such as birds and terrestrial mammals, whereas it has been shown to be more responsive than abundance to AES intervention in some studies on invertebrates (Section 2.3.5).

Figure 26 Focus of pollinator and butterfly sampling in focal and surrounding 1 km squares. Solid red lines mark 0.5 km sampling extent within surrounding squares. Dotted red lines denote 3 × 3 km landscape sampling unit.

4.1.1.5 Butterfly and bumblebee transect protocol (based on Pollard transect methodology)

Data will be collected for all butterfly and bumblebee species, with each transect section walked once to record butterflies, a second time to record bumblebees, and a third time to assess floral resources. Transect sections will be walked at an even pace and only the butterflies and bumblebees observed either visiting flowers or in flight within a 5 m box around the surveyor recorded (up to 5 m in front, 5 m above ground and 2.5 m either side; see Figure 27). Anything which is flying further ahead or otherwise outside of the box will not be recorded. Transects will be walked between 9.30 and 16.30 hours, when the temperature is at a minimum of 13 °C when
it is sunny with 50% cloud cover or less, and 18 °C or more when cloudy. A net may be used to catch individuals that are difficult to identify but it is assumed that, as with the UKBMS / WCBS, the majority of butterfly species will be identifiable on the wing. Bumblebees will also be identified to species in the field (with the exclusion of workers of two very similar species) and total abundance of bumblebees recorded. Once the transect section has been walked in one direction, it will be walked again in the opposite direction to assess floral resources using an index that has been used in previous research (Carvell et al. 2016).

**Figure 27** Moving box sampling approach for Pollard transects, used by WCBS and UKBMS (taken from (van Swaay et al. 2012)).

| 4.1.1.6 Pan trap protocol for bumblebees, solitary bees and hoverflies |

**Figure 28** Pan trap station for sampling pollinating insects, used in NPMF project (Carvell et al. 2016)

At one end of six of the ten 200m transect sections a pan trap station will be placed to sample solitary bees, hoverflies and bumblebees (as well as other flies, beetles and insects referred to as ‘by-catch’ but with potential for future analysis). Each station will consist of three pans or bowls, one sprayed each of UV-reflective white, yellow and blue, as specified in Carvell et al. (2016).
All pan traps in a sampling square will be set out prior to starting the transect walks and will be collected 24 hours later in the same order in which they were set out. Invertebrate samples will be stored in alcohol, sorted to separate pollinator groups, and all bumblebee, solitary bee and hoverfly species identified subsequently in the laboratory.

4.1.2 Monitoring moths

4.1.2.1 Methods for sampling moths

The studies reviewed used two methods for sampling moths: 1) light traps and 2) transects. Light traps were by far the most frequently used method of collecting moths (e.g. Littlewood 2008; Merckx et al. 2009a; Merckx et al. 2009b; Taylor & Morecroft 2009; Merckx et al. 2010a; Merckx et al. 2010b; Fuentes-Montemayor et al. 2011; Heard et al. 2012; Merckx et al. 2012b; Alison et al. 2016). Moths are mainly nocturnal fliers, so are effectively sampled during the evening / night. Light traps provide a light source dominated by the UV end of the light spectrum, which acts an attractant to nocturnally flying moths, which are then trapped inside. Like any sampling method, light traps may be biased towards different taxonomic groups, and their effectiveness partly depends on local conditions. Merckx and Slade (2014) used mark-release-recapture of individuals from three common families of macro-moths, Erebidae, Geometridae, and Noctuidae, to show that individuals from different families were attracted to light sources from different maximum distances (corresponding to a 5% recapture rate), 23m, 27m and 10m effectively. Further, they predict that when flying close to a trap (0-1m) a maximum of 55% of Erebidae would be attracted to a trap, 15% of Geometridae and 10% of Noctuidae. As such, light traps are likely to be biased in both the size of the locality from which they attracted moths from different families, and with regard to the proportion of those moths that encounter the trap that will be caught. However, as long as traps are located in the same position on each occasion such biases should remain consistent, so light traps are effective in detecting changes in abundance and species richness over time.

Three studies reviewed used observational transects to record moth abundance, but two of these only reported aggregate abundance with other insect groups as a response (Ekroos et al. 2008; Franzén & Nilsson 2008) and the other focused only on the aggregated abundance of species of the day flying burnet moths (Zygaena spp.), which due to their striking black and red marking are easily distinguishable in the field. Moth transects differ from those proposed for pollinating insect and butterflies (Section 4.1.1), in that sweep netting of the vegetation is used to collect moths for identification and therefore. Given the amount of the disturbance this may cause and different search pattern, it is not suitable to be conducted at the same time as the pollinator and butterfly transects. Two studies also used sweep-netting (running a net through the vegetation), by sampling a set area over which they conducted a fixed number of sweeps, however both studies reported results only for aggregated abundance (De Snoo & Leeuw 1996; Hahn et al. 2015).

Macgregor et al. (2016) recently used a night transect method for moths, adapting the protocol of Birkinshaw and Thomas (1999) by walking 25m transects at night, aided by a red filtered floodlight and head torch, catching moths using a hand net at 1m intervals for 1 minute each.
Further, they used overhead search light activity surveys, counting the number of moths seen passing through an upwards pointing torch beam, to estimate total moth abundance. They conducted light trap sampling concurrently and found that the compositions of each sample based on either species or family were significantly dissimilar. As such, this may show promise as a method to complement light traps, much like transects and pan traps can be used complementarily to sample pollinators. However, the logistical difficulties of using such a method at night across a large number of sampling units are likely to be challenging.

Moths are frequently split into two groups, macro- and micro-moths. The majority of AES-related studies on moths have focused on macro-moth species, which tend to have greater mobility (e.g. Merckx et al. 2009a; Merckx et al. 2010a; Merckx et al. 2012b; Alison et al. 2016), with just three studies also recording micro-moths to species level (Fuentes-Montemayor et al. 2011; Heard et al. 2012; Staley et al. 2016) and one assessing aggregated micro-moth abundance (Littlewood 2008). It is possible to identify the majority of macro-moth species collected with light traps in situ in the field, and then let them go. In contrast, many micro-moth species need to be identified in the laboratory, due to a requirement for microscopes and in many cases dissection. Macro-moths are therefore cheaper to sample, given a reasonable level of proficiency in moth taxonomy, while more potential recorders are also familiar with macro-moths than micro-moths. There is some limited evidence that at the local, farm scale micro-moth species may be more responsive to AES management than macro-moth species (Heard et al. 2012), although a second study found that both macro- and micro-moth species responded to AES interventions in slightly different ways (Fuentes-Montemayor et al. 2011). Given the likely limits on resources available for this sampling, the protocol below is based on macro-moth sampling and identification in the field, while recognizing that this does limit sampling to only part of the moth assemblages present on farmland. An additional module to include laboratory work for micro-moth identification could be included if required.

4.1.2.2 Protocol for sampling moths

Given the available evidence and their ubiquity in studies assessing AES effects on moths, we propose to use portable Heath light traps to sample macro-moths, with sealed lead acid batteries for power.

Each sampling unit will be sampled at a minimum twice, once in late May/June and once in July/August. There is high species turnover for moths, with many species only present as adults for 2 – 3 weeks. To increase the likelihood that the same moth assemblage is being assessed across sampling units, it is important that light traps are moved on to successive sampling units after a single night to allow as many units as possible to be covered in a given 2 – 3 weeks. If possible more than one sampling unit should be surveyed concurrently in a single night, although logistical constraints may prevent this.

Sampling will only occur under minimum defined weather conditions. For example, forecast criteria of a minimum night temperature (10 °C), a maximum wind speed (20 km/h) and maximum precipitation risk (50%) will be applied, following previous studies sampling moths in agricultural habitats (e.g. Merckx et al. 2009a; Merckx et al. 2010a).
4.1.2.3 Sampling moths within focal 1 km squares

Six portable Heath light traps will be placed in each focal 1km square, in the same position as the six pan traps used to sample pollinators (Section 4.11, Figures 24 & 25). Moths frequently use linear features to move between semi-natural habitat in agricultural habitats, as well as providing shelter (Slade et al. 2013), so this is likely to maximize the number of moths sampled by the light traps (Alison et al. 2016). Some degree of co-location of sampling for pollinators, butterfly, moths and plants, as proposed, will allow data collected from the same locations within each 1 km square to be compared across these taxa.

Light traps will be set in the evening, on the west-facing side of the linear feature to minimize morning sunshine, which can increase moth activity and cause them to leave the light trap. The entrances to all six light traps in a single 1 km square will be plugged the following morning as close to dawn as possible to prevent the moths escaping; each light trap will then be visited again to identify the macro-moths. All macro-moth individuals will be identified in the field (excluding a few cryptic species that will be collected and identified in the laboratory). Photographs will be taken where necessary to aid or confirm identification. Moths will be released at the location from which they were caught.

4.1.2.4 Sampling moths in surrounding eight 1km squares

Moth sampling in the wider landscape sampling units will focus on subsampling the same areas as the pollinator, butterfly and plant protocols, i.e. half of each of the four adjacent 1 km squares (Figure 26). Three Heath light traps will be placed in each of these four 0.5 × 1 km rectangles, in approximately the same positions as the pan trap stations. By keeping the sampling density the same in the sub-sampled surrounding squares, this will allow estimates of moth species richness to be directly comparable between the focal and outer squares, as discussed in the pollinator and butterfly protocol above (Section 4.1.1.4). Where possible, sampling in the focal 1km squares and surrounding squares will be conducted concurrently, or on consecutive nights.
4.1.3 Monitoring birds

It is important to measure bird numbers and locations both in focal squares and at the landscape scale, so to sample at both scales. The critical response variable will be breeding bird numbers, either counts of individuals or counts of territories (often via numbers of singing males as a proxy). Winter monitoring is also desirable to assess the performance of options aiming to provide winter food (see Section 4.1.3.1 below). For long-term monitoring of CS effects, it will be important to use a protocol with consistent, repeatable survey intensity and coverage, which is as independent of surveyor identity as possible (allowing for changes in surveyor over time without compromising inference). In addition, to compare focal squares with the surrounding landscape, the coverage needs to be similar in intensity per unit area sampled.

The BTO/JNCC/RSPB Breeding Bird Survey (BBS) methods are designed to be quick and simple, for use by volunteer surveyors, and to provide inference at the regional scale and above. However, more intensive surveying is required to measure local abundance and habitat selection accurately, such that bird numbers, both across the entire survey unit and within individual habitat patches, can be reliably linked to habitat features, management or survey records for other taxa. All volunteer surveys are also subject to a degree of uncertainty, because coverage of any specific location cannot be guaranteed. Therefore, the targeted monitoring of CS management impacts and counterfactual areas at multiple spatial scales cannot be achieved by exploiting ongoing BBS survey effort, while re-targeting survey squares with respect to CS coverage would both compromise the existing survey design and be unsustainable because of unpredictable turnover in square coverage by volunteers. However, BBS data can provide important contextual information on regional abundance and population trends, as well as allowing scaling up from focal NCAs to national patterns. To facilitate this, new, bespoke survey results that can readily be linked to those of BBS would be most valuable.

The survey method proposed here operates within the broader sampling design of a focal 1km square and the surrounding “landscape sampling unit” of eight 1km squares, making a total survey area of 9km² per sample site. Surveys would be repeated four times between late March and early July, each year. The BBS method features two visits to 1km squares, with 2km of transect walked per 1km square. A fixed length of transect through sample squares is also desirable for CS evaluation, to standardize survey effort between visits and years (as opposed to a freer, “roaming” design), but more complete coverage of the focal square than in BBS would provide more monitoring power and more visits are needed to reduce the stochasticity in estimates of local abundance. Transects are considered to sample up to 100m on either side of the line followed (in farmland), although detection and coverage will clearly be more intensive closer to the transect itself. This transect should intersect, cross or be adjacent to a representative sample of the bird-relevant AES management in the focal square. However, transects must also be repeatable; for example, if a given field is under a rotational option, it might be desirable for the transect to cross it, but if the field is likely to be under oilseed rape or to hold a bull in a future year, such a transect would then be impassable, so should be avoided. Overall, 3km of transect within a 1km square, walked four times each year, will provide repeatable data that are close to complete counts for 60% of the surface area of each focal square, assuming detection of the birds present up to 100m from the transect line. Limits to accessibility (which will mostly be along linear features, even where access permissions are available) are likely to restrict possible
coverage of many squares in practice, such as by making following longer lengths of transect impossible with risking double-counting of individual birds. If areas of a square are impassable, such as lake surface, active quarry or military firing range, the transect length will be reduced proportionally with the available survey area. These data will also be “degradable” to be equivalent to BBS data, for comparison, by sub-sampling transect lengths and visits. The area surrounding the focal square cannot feasibly be surveyed with the same intensity, but a consistent protocol needs to be used to deliver comparable abundance data. Hence, transect routes of 8km in total crossing the eight 1km squares surrounding the focal 1km square, each followed four times, would provide appropriate information. These transects should sample the landscape in the entire area concerned as representatively as possible (i.e. not every component square necessarily needs to be covered with the same length of transect or effort) and be repeatable in the long term, in the same way as the transects in the focal square. The complete transect design is illustrated shown as a set of possible transects for a real 1km square plus surrounding eight 1km squares in Figure 30. Pragmatically, these transects would best be connected or near-connected to facilitate efficient coverage in practice.

Field surveyors should work using large-scale maps of the survey area, with transect routes marked on them, for recording the locations of all birds detected. Maps produced using unlabelled OS MasterMap polygons work well for this. All birds seen and heard should be recorded, allowing retrospective assignment to distance bands for comparison to BBS results. Mammals seen and heard should also be recorded (see mammal protocol). The mapped bird and mammal records should be digitized to facilitate this process and to allow assignment of records to habitat and management parcels. The output data would then comprise species-specific counts per unit area (i.e. densities) in the focal square and the surrounding landscape, and counts per
habitat parcel. These will be visit-specific, so can be analysed in this form (using a repeated-measures approach, see below), or combined across visits to represent likely territories or to take means or maxima. Square-level data are valuable as measures of local abundance for measuring changes over time and the influence of CS management. In addition, habitat-patch-level data reveal the selection or avoidance of areas under scheme management, providing both a test of potential efficacy in resource provision for target species and data to inform analyses of causes and mechanisms behind changes at the square level. The latter would involve analysing population change with respect to the use of CS-managed patches, reflecting their efficacy in practice, to add to analyses of associations with quantities of management (option area or length).

It is important to note that the survey approach described above would be effective for the detection of management effects on species that are widespread across the selected sample of focal squares, but unlikely to detect effects on very localized species. Evaluation power for colonial species will also be limited because it is less likely that aggregations of a given species will be found in areas with a range of AES management quantities than is the case for more dispersed species. Nocturnal species (i.e. owls) would also need additional bespoke monitoring approaches.

4.1.3.1 Winter bird monitoring module

Policy targets and ongoing annual bird monitoring consider only breeding bird populations in terrestrial habitats in Britain. However, some CS management aims to benefit birds in winter. Some effects of winter management can be assessed by analyses of breeding population change, because populations are resident and locally breeding birds can meaningfully be linked to local winter habitats. However, this ignores potential benefits for winter immigrants, such as thrushes or finches, and may be insensitive for more mobile “resident” species (because local winterers disperse too far to breed). Winter monitoring is needed to fill this gap and to provide habitat-patch-association data relevant to assessing the likely efficacy of winter options and to allow analyses considering mechanisms of action. Late winter is also the best time of year to record breeding populations of mammals such as hares (see mammal protocol). Nevertheless, a coherent programme of monitoring and evaluation could be conducted using only breeding population monitoring, so winter surveys are presented here as an additional module.

Winter surveys would aim to measure the bird use of relevant option types (probably over-winter stubble management, wild bird seed crops and hedgerows), as well as estimating background over-winter abundance. Survey routes are more difficult to standardize than breeding season surveys because many of the options will be rotational and different routes will be required each year. Further, whole-field habitats in winter need to be surveyed using whole-area search counts, because birds are commonly in flocks within fields and cannot be surveyed efficiently using transects or point counts. Winter surveys would therefore follow a fixed route within and between winters, but would not need to follow the same route as the breeding season surveys. In addition, field-specific, whole-area-searches of fields under winter seed options (stubble management and wild bird seed crops) would be conducted, either as part of the transect or separately. Surveys would be conducted monthly (November-March) and would follow a similar mapping approach to the breeding season ones. The transect element would aim to cover a total
transect length of 9km, evenly spread through the 9km² survey area and sampling the habitat present broadly representatively. Transect routes would take in lengths of relevant hedgerow options and out-of-scheme hedgerows.

The output data would comprise counts of birds in CS option patches, for comparison with the background numbers in the surrounding landscape from the transects. These data would provide both a test of likely local option efficacy and indices of habitat use to inform the analyses of effects on breeding abundance. Monthly data could be averaged or considered by month, because the local population context and available resources will vary through the winter, so effects might be most detectable, or most important, in certain months (e.g. when weather conditions are most harsh or when ambient food resources are most depleted).

4.1.4 Monitoring terrestrial mammals

The potential to monitor AES effects on mammals is limited and complicated by the specificity of the survey approaches required for different species. Thus, unlike with birds, for which a single protocol can suffice for the vast majority of species, a wide range of sampling techniques would need to be applied to monitor all mammals in Britain. In addition, many approaches, such as those based on field signs, are likely to be effective for detecting species’ presences, but rather insensitive to (or unreliable for) changes in abundance, when the latter are more likely to respond to management rapidly in practice. However, there are rather few mammals that are targets for AES management and therefore priorities for the monitoring of its effects. Three species were identified as potential priorities (on the basis of being potentially responsive to AES rather than as conservation status priorities) in the project workshop: Brown hare, Field vole and Harvest mouse. In addition, research has shown that hedgehogs can use and benefit from key AES management options (Hof & Bright 2010), and a monitoring technique has been proposed (Yarnell et al. 2014). The hare approach would be a component of the bird protocol, while the other protocols represent independent modules.

4.1.4.1 Brown hare

Brown hare, as a large, diurnal species of open field habitats, is well-suited to monitoring with visual surveys. Moreover, it is readily recorded during bird surveys, because the basic search image and scanning approach is compatible with that used in practice for large ground birds, such as partridges. Weaknesses with linking hare and bird surveys are that (i) the timing of typical breeding bird surveys in spring comes after the hare breeding season, such that counts of “adults” could be artificially inflated by the inclusion of juveniles, and (ii) that hares are more accurately crepuscular than diurnal, so counts made later in the day and later in the season may be under-estimates because lower activity reduces detectability, especially in taller vegetation. However, a combination of hare monitoring during breeding and winter bird surveys, considering time of day (e.g. via the up-weighting of observations from earlier in the day) should be effective in producing reliable data. In addition, as with the bird survey protocol, this approach would allow data matching to those being collected under the BBS to facilitate scaling up.
4.1.4.2 Field vole and Harvest mouse

After conducting several trial projects (Sibbald et al. 2006; Poulton & Turner 2009), The Mammal Society has proposed a monitoring scheme for small mammals, with a protocol that is designed for volunteers and the large-scale detection of change in distributions (although note that the scheme is not currently actively being promoted). Hence, survey units are short (100m) transects in suitable habitats within 2 × 2km tetrads. This gross sampling design is not appropriate for monitoring AES responses because the 2 ×2 km sampling units are very large, and much larger than patches of AES management would be. However, the specific transect component could be used, therefore providing a link to future national volunteer scheme data for scaling up, although with a greater sampling intensity within focal monitoring sites. Therefore, a possible survey design would involve 100 m transects for each species, up to a maximum of four in each focal and surrounding 1km square, within areas of suitable habitat, including any patches of relevant AES-managed habitat. In practice, the habitats favoured by these species will be sufficiently rare that many fewer than the theoretical maximum of 90 transects per 3×3 km sample unit would be identified as requiring coverage. Many NCAs will also be extra-limital for harvest mouse, in which case surveys for this species would not be necessary.

Both Field vole and Harvest mouse are recommended to be sampled via field sign transects, the former in autumn and the latter in December, with just a single visit per site per year. These surveys are not, therefore, likely to combine readily with those for other taxa, so they will entail additional costs, although late autumn and winter sampling could be combined with winter bird surveys. At the transect level, field surveys would follow the Mammal Society protocols. Harvest mouse nest transects comprise ten, contiguous, 10 × 10m squares covering a total area of approximately 1,000 m$^2$. Nests are counted in each square, providing ten integer counts from zero to $N$. This method requires only a single visit, with on-site time probably in the range of 1-2 hours.

Field vole sign transects would be overlaid on harvest mouse nest transects, if the habitat is appropriate, or be independent. They comprise ten 1 × 1m quadrats at 10m intervals along a 100m transect (leaving 5m at either end). Within each quadrat the presence of runways (worn paths weaving through the grass stems with evidence of chewed-off grass stems), latrines (collections of green/dark green faeces) and feeding signs (clippings of bitten-off grass stems and leaves often left in a criss-cross pattern) are recorded separately. These measures each inform more reliably in different habitats. This method also only requires a single visit, probably taking slightly longer than harvest mouse nest transects.

Poulton & Turner (2009) investigated the power of the Harvest mouse method and found that intensive, frequent visits were needed to detect short-term fluctuations in abundance. Long-term variations are of more interest for AES monitoring, so the concise approach is more likely to be sufficient. However, power analyses suggested that many thousands of sites would be required to detect changes in patterns of presence and absence over a five-year period, and that nest count data, while delivering more power to detect changes in abundance, still required a sample of 2500 to detect a 30% decline over five years (Poulton & Turner 2009). No analyses of the power to detect effects of management have been conducted, but this suggests that the power provided
from this monitoring approach may be low. This should be considered before monitoring for these small mammals is rolled out in a new programme.

4.1.4.3 Dormouse

Dormouse is one of the few mammal species referenced specifically in AES targeting in England, with respect to “high-value hedges” (e.g. the HLS handbook). The published survey approaches for this species involve placing artificial nest sites (tubes or boxes) in suitable habitat and checking subsequent occupancy as an assay for presence (Chanin & Goubert 2011). Densities used have been 5 per 100m of linear habitat (Bright & MacPherson 2002; Chanin & Woods 2003), a loose grid of 10-30m depending on habitat density in larger habitat patches (Chanin & Woods 2003) or a minimum spacing of 70m (approximately one per hectare; Mortelliti et al. 2011), all of which provided sufficient power to detect variations in dormouse presence in the habitats and regions considered. Bright & McPherson (2002) set 200×60mm tubes in March-April and checked them the following June and September-October; Mortelliti et al. (2011) set boxes at a similar time and checked them in “spring-summer” and “autumn-winter”. Chanin & Woods (2003) provide more comprehensive recommendations and suggest that tubes are installed no later than April and finally checked no earlier than October, with early March to the end of November being the ideal period and the end of July to the end of September being the minimum. Chanin & Goubert (2011) compared the performance of boxes and tubes and found that boxes were more likely to be occupied and therefore more sensitive as an assay of presence, and also superior for monitoring breeding success. However, they are also much more expensive, meaning that tubes are a better option for intensive surveys over large areas.

A tube-based protocol could readily be applied in sample units in NCAs within the range of the dormouse in England. The specific design would involve 100m sections of suitable (hedgerow) habitat within the focal and surrounding 1km squares, including within all hedgerows managed with relevant AES options. Following Bright & McPherson (2002), five nest tubes would be placed, evenly spaced, in each 100m section of hedgerow, with 5-10 100m sections selected in each 1km square, dependent upon the landscape and availability of potentially suitable habitat in the landscape. Population monitoring would consist of nest counts in the tubes in the autumn. This protocol would entail additional costs for the field visits in autumn, although the tube-setting could potentially be combined with early spring bird or plant surveys, or with a visit to set refugia for reptile sampling. Tube materials would also entail a capital cost. A caveat to the use of this approach is that there is no evidence as to its power to detect effects of management. The studies conducted to date suggest that effects on dormouse presence could be detected, but it is unknown what degree of change in abundance would be detectable and all studies to date have considered only detection of presence. A further consideration is that, while no licence is required to set dormouse tubes, one is needed to handle or to disturb the species, so any positive records require any subsequent checking of tubes or continuation of surveying to be licensed (Chanin & Woods 2003). Suitable training courses are, however, readily available from the Mammal Society, so it would be important for surveyors to attend such courses and to have experience of dormouse handling, to obtain the necessary skills required for handling licences to be granted.
4.1.4.4 Hedgehog

Hedgehogs will not be detected reliably by the visual surveys conducted primarily for birds. However, an approach based on baited footprint tunnels has been developed and tested by Yarnell et al. (2014), which could be applied as an additional survey module. Ten baited tunnels would be placed within suitable habitat in the target 1km square checked each morning over four days during the spring and summer (April-September). Three tunnels would be similarly placed and monitored within each of the surrounding eight 1km squares. The lack of sensitivity to survey timing during the summer (Yarnell et al. 2014) means that this method could be applied flexibly and potentially combined with other taxa surveys to reduce costs. The principal limitation of this method is that it has only been trialled as a method for measuring (changes in) presence, for which it can readily detect change or variation of 25%. Further piloting would be required to investigate possible interpretation in respect of variation in abundance (conflict with variation in activity and breeding success seems likely). A change in occupancy of 25% may also be difficult to achieve (or to prevent) with AES management, so this method may not provide useful evaluation power for AESs. Note, however, that there has only been one survey of hedgehog population sizes anywhere other than Uist over at least the last 20 years (Parrott et al. 2014), so any broad-scale quantitative monitoring that can be conducted would be extremely valuable in its own right, given the likely national decline of the species (L. Kubasiewicz, pers. comm.).

4.1.5 Monitoring bats

The survey approach would use full spectrum bat detector technology, where detectors are left out to trigger automatically all night and to capture bat calls at their original frequency, retaining more detail of the call than other detector types. When used in conjunction with call identification software and validation (see Newson et al. 2015)), this approach offers great potential for transforming large scale bat monitoring in the UK, and for addressing the specific question in relation to AES effects, including woodlands and woodland bat species (see Section 2.5.5). Acoustic identification of some bat species, in particular Myotis species, is challenging but rapidly improving and by retaining recordings there is the potential to re-analyse them as processes improve. Using passive detectors in this way, there is the potential to provide representative acoustic monitoring of bat species distribution and activity as a measure of relative abundance. This approach can generate a large volume of recordings per night as a measure of relative abundance, but the number of recordings can be highly variable depending on nightly weather conditions, local habitat and use of particular features in the landscape, for example proximity to a roost. To address sampling noise, objectivity, replication and habitat representation is required. Static detector recordings are likely to provide a useful and objective standardised measure of relative abundance, and the only way of getting landscape scale measures of relative abundance for bats. In addition, there are promising analytical methods (see 2.5.5 above) being investigated to help understand the relationship between relative abundance and real numbers. It is unlikely that alternative sampling methods, such as collecting occurrence-only data through walked transects, would provide these qualities.
The specific field protocol will involve three detectors deployed simultaneously within a 3 × 3km square (i.e., the core 1 km plus outer spill-over area), as per Figure 31, to cover: (a) key AES options, (b) the 1-km core square and (c) spill-over squares (see 4.1.5.1 for further details). The detectors would be deployed for at least three consecutive survey nights in one location, before being moved to new sampling locations within the same 3 × 3 km square, before then being moved on the seventh day to a new 3 × 3 km square (for a further 6 day deployment). Thus, three detectors are shared and rotated between four 3 × 3 squares per month. The multi-night deployment per location is important in order to help to average out variation in bat activity due to weather, but the total number of recording nights may depend on logistical considerations in practice. Sampling between May and the end of September would provide seasonal representation for breeding abundance (May-July) and post-dispersal habitat use (August to September). Moving the detectors within a 3 × 3 square will help to control for habitat heterogeneity, which may influence bat activity on a given night. Newson et al. (2016), showed that by using this approach some species are recorded more widely post-breeding and additional exploratory work with these data has also demonstrated seasonal habitat use. For example, greater use of farmland habitat by the Brown long-eared bat later in the season. This is in support of Stephanie's Murphy's radio-tracking work on this species limited to a small number of radio-tracked individuals, and of course radio-tracking would also struggle to detect wide ranging individuals moving beyond the farm level across landscapes. The detectors should not be deployed within 100m of moth traps.

For reasons of detection efficacy, the detectors would best be deployed in a stratified way following habitat appraisals for the survey squares. This will avoid inevitably low detectability in open expansive areas of land. The metric would then be ‘bat activity’ as measure of relative abundance that is comparable within the three zones (AES, 1km focal squares, and 3 × 3km buffer), and over time.

![Figure 31 Proposed sampling framework for bat and amphibian/reptile monitoring](image-url)
4.1.5.1 Potential sampling framework for bats

The general principle is to synchronise sampling for each taxa across three zones as follows:

1. Within the AES option or AES concentration ‘zone’, beginning at random sample point S1.
2. The inner 1-km focal zone beginning at random mobile sample point S2.
3. The outer zone (3km x 3km), beginning at random mobile sample point S3.

As bat activity is affected by the localised influence of habitat and particular landscape features it is imperative to install duplicate if not triplicate simultaneous sampling within the zone S1, S2 & S3 but also to move bat detectors across the zones to control for habitat heterogeneity and localised influences such as weather conditions on activity patterns. Importantly, as bats typically utilise structural components of landscapes, sampling cannot be random but must include an appropriate level of habitat stratification (targeting AES options that are expected to be beneficial to bats) and careful pairing of sample locations to increase the probability of bats being detected, to maximise the power of the method to detect the effects of AES on the relative abundance of bats and to improve the probability of detecting change. Otherwise, the intensity of repetition and the degree to which sample points 1, 2 and 3 are repeated within the three zones, even with stratification, requires further investigation, probably through pilot testing. The framework would be repeated within high and low-density AES areas, along the gradients sampled in each NCA.

4.1.6 Monitoring amphibians and reptiles

Measuring population abundance in a meaningful and comparable way is difficult for most species of reptile and amphibian. Difficulties and variability in detection present sampling issues while lack of understanding of population dynamics, especially for species that show large fluctuations in population sizes, make interpretation problematic. The general rationale for amphibians and reptiles would be to use habitat proxies to determine population viability and status while using field surveys to confirm presence of the species. While expert opinion and current understanding of the ecology of these species will allow this approach to be initiated quickly, this would however require intensive methods to be developed early on to determine the relationship or correlation between habitat quality and abundance (ideally counts over a unit time element to account for effort). This would inform an extensive method using habitat suitability criteria (albeit with ground-truthing from time to time) to deploy as a method for surveying larger scales such as landscapes, potentially by using remote methods of detection such as using LIDAR or drone technology to assess habitat viability per unit area (e.g. per 1km square).

The methods should cover (i) widespread, politically important species (Great crested newt) and (ii) rarer, more habitat specific species, though these species are often highly associated with designated sites (many of which are operating AES). Methods should be developed to allow assessment of the other S41 priority species of reptile and amphibian. It is likely that the rarer,
more habitat specialist species will require a more detailed sampling framework than the more widespread species that occur over a range of habitat types.

Ideally, within a 1 km square, an overall monitoring strategy would use generalised techniques, such as habitat assessments (e.g. via ‘expert opinion’ and/or remote sensing to measure vegetation age, structure, connectivity and permanence), validated, via field sampling, using a simple ‘presence’ metric (or possibly a ‘confirmed breeding metric’), ideally with stratification by gross land cover, to ensure representation. Such site level assessments could, over time, be supplemented by species distribution modelling based on independent data sets, which could allow more sophisticated analysis of the probability of occurrence based on habitat condition/proximity of records, although in the initial stages we would advocate a focus on field based assessments. Logically, the appeal of this approach, accepting fine tuning via ground-truthing, would be that possible ‘presence metrics’ could be applied to multiple taxa with the same baseline environmental data set. This approach could be overlaid on AES areas, with AES acting as a subset of ‘all habitats’. GIS analysis (and species distribution modelling) would allow the effective contribution of each measure to be evaluated and, where prescriptions are wrong, resulting in a decline in habitat quality then the ‘species metric’ would also show a decline.

The breeding behaviour of amphibians allows a pond based approach and, through this, some opportunities for gaining abundance measures. However, it is unrealistic to expect to get reliable abundance measure for reptiles in the wider countryside without considerable effort, as detection rates are likely to be low and erratic. Nonetheless, in any sampling regime quantitative data should be collected to help interpretation. However, when looking at habitat as a proxy, it is important to use metrics that would be expected to impact on population parameters (population ‘size’), such as habitat extent, habitat quality, range and future prospects using a measure of ‘viability’ (e.g. four occupied ponds with Habitat Suitability Index (HSI) score >0.7) or via a minimum area of high quality habitat derived through predictive modelling. The underlying ‘rules’ for determining such ‘viability’ proxy measures would need to be developed over time though indicative values could be provided relatively quickly for the purposes of developing the methodology. For example, while some species, such as Grass snakes, can occupy a number of habitats other reptile species may show a much more restricted and clumped distribution, and generally habitat connectivity will be important to link isolated features and maintain population viability. Key species measures, where more detail is required, would be made via a sampling regime. Volunteer effort is unlikely to be sufficient for intensive surveys based around AES, although a national scheme could achieve significant uptake if designed and implemented appropriately.

Two parallel sampling regimes are advocated.

1. Deep and narrow: an intensive sampling regime within a restricted range of sites will allow the greatest opportunity to measure abundance and population change, and reducing reliance on preconceptions and resulting biases. However the focus on few sites will risk being unrepresentative of sampling units and the possibility of untypical results.
2. Broad and shallow: this would look at a larger number of sites over a broader geographic area, with a focus of habitat quality metrics and detection of presence. This will offer the benefit of allowing comparisons over a wider range of sites and so is likely to ensure a more
representative range of habitat conditions and outcomes. This is particularly important as the clumped nature of the occurrence of many species around habitat features may lead to under or over representation of species in some areas. More extensive sampling may allow a clearer picture of the benefits of different schemes.

Both approaches should be run in tandem, although it is likely that there will be an increasing shift from the ‘deep and narrow’ to ‘broad and shallow’ as ecological correlates are improved within the sampling regime, environmental data (e.g. from remote sensing) becomes more detailed and more readily available and as species distribution modelling becomes more accurate.

Importantly, a ‘population’ measure or a landscape-level response to AES would have to be supported by an effective programme of ground-truthing and ‘validation’ of animal presence at distances from the AES concentrations. Important characteristics are as follows:

1. There is low dispersal potential for most amphibians and reptiles of < 1km.
2. Species may be ‘clumped’ and easily missed by random searching.
3. The sample unit would need testing (e.g. each 1km with a 5km radius)
4. For comparative studies (e.g. AES vs non-AES), it is likely that fewer, more intensively studied sites that provide a real metric of population abundance will be most useful.

4.1.6.1 Potential sampling framework for reptiles

Sewell et al. (2013) note that the most effective methods for assessing relative abundance counts or densities of reptiles is to combine (a) refugia sampling (or artificial cover objects) with (b) transect walks (i.e. visual observation of active animals along transects), ideally targeting suitable habitats and habitat features that maximise the opportunities for finding the species. Any attempts to provide abundance/relative abundance measures will need multiple visits and the reliability of these measures is hard to ascertain due to the highly variable detectability of reptile species.

For this reason, nominally, the diagonal transect walk has been added to Figure 31, as reptile transects (‘RT’) 1 and 2, with RT2 moving randomly between outer zone squares between visits. However, it may be that a different arrangement of transect walks would be recommended by experts. Surveys can be undertaken throughout the spring and summer months, but are usually focused around April and May for adults and September for young (Sewell et al. 2013) when optimal surveying conditions are likely to be available for longer. In the wider countryside, this method might only apply to the more widespread species: Grass snake, Common lizard and Slow worm (possibly Adder), and may be applicable to the rarer species (Sand lizard and Smooth snake) on protected sites only (although often a more intensive sampling regime is recommended for the rarer species focusing on their specific habitat preferences). For Adders the most effective sampling regime is counting when they emerge early in spring at known hibernation sites, but this is unlikely to be widely applicable for comparative studies relating to AES schemes. In such sites rapid drop-off of species presence is expected outside the AES-managed area, due to abrupt changes in land-use and habitat (so therefore little chance of a detectable landscape effect of AES).
Reptile detectability is unpredictable and is affected by habitat suitability. For some species, notably lizards, this results in clumped distributions around optimal habitat or key features; more mobile species (i.e. snakes) may show predictable seasonal movements between areas and may occur in a broader range of habitats, but overall are likely to be more prevalent in good habitat. Sampling for reptiles tends to focus on selecting areas where a priori expert judgement indicates the greatest chances of success, with such bias being accepted as necessary to achieve adequate sample sizes. However ‘optimal habitats’ are often those providing the greatest variety of habitat structures and as such will offer different detection rates to other habitats.

Reptile surveys can be standardised, using a fixed number of refugia (usually c. 10), roughly consistent survey areas (e.g. <25 ha) and fixed survey periods, usually aiming for 1 - 3 hours. However, often a bespoke survey design is needed to address specific questions and especially to accommodate the need for increased effort where detectability is predicted to be particularly low. Where it is considered likely that sample effort or detectability is likely to be insufficient to provide meaningful comparison, greater emphasis will be placed on habitat quality assessments and with a lower demands being placed on species data – and generally looking solely at presence, extent of occupancy or some measure of viability (evidence of breeding, sightings of >1 individual/ age class).

Following Sewell et al. (2013) with intensive surveys, using 30 refuges and 7-11 visits should reveal the species that are present and provide measure of relative abundance. However, away from key features it is likely that low levels of detectability may prove problematic for analysis, although this is likely to require more replicates of the different regimes rather than increasing sampling within each one. Thus we would advocate that habitat assessments, looking at both ‘quality’ and extent, should form an integral part of any survey regime.

4.1.6.2 Amphibians

For counts at the landscape level, bottle trapping, netting and torching are favoured methods, with visual surveys, during March (frogs and toads), and mid-April - May (newts), using six survey visits. Environmental DNA sampling is becoming increasingly popular and available for presence/absence surveys for great crested newts, but currently is not suitable for assessing relative abundance (and has not been developed for reliable field assessment for other species in the UK as yet, Biggs et al. 2015). The methods pertain to ponds and ditches, so a very different protocol or framework is needed from that in Figure 1, targeting wet habitats only. Repeat surveys can provide count data that in turn may allow assessment of relative abundance. However, such counts can be unpredictable and will be heavily affected by visibility and habitat parameters. Bottle trapping and netting can allow greater standardisation, but are more intrusive, significantly more labour-intensive and require greater levels of skill. Spawn counts can provide a useful additional metric for Common frog and Natterjack toad; for Common toad and newts, spawn is most often used as a measure of presence. For all methods there will be temporal difference in peak abundance for the different species and any attempt to gain meaningful counts will need to involve multiple visits around the peak times for each of the target species.
A sampling framework for amphibians would usually involve looking at all ponds in the sample area, looking at presence and for great crested newts where there is an established method, undertaking an HSI score. Survey protocols are determined by the likely species and by habitat condition with around six visits being focused around peak count periods for each of the target species, if quantitative data are needed. The choice of methods will be influenced by detectability in the waterbody.

Assessment of the use of terrestrial features, such as buffer strips, can involve terrestrial surveys, using cover objects. This is intensive and has a generally low detection rate, but might be the only viable option for assessing the value of specific features versus gaining some overall assessment of population levels within a landscape as ponds are likely to attract amphibians from 50-500m (and sometimes greater distances). Sampling methods have been developed for comparison of buffer strip features (Salazar et al. 2016) in which 100m sampling transects were established and replicated on features and between different farms. On each transect, at 20m intervals, five sets of triplicates of ‘cover objects’ using different materials (roofing felt, onduline, carpet) are set to offer a range of different conditions to attract terrestrial amphibians and reptiles. Each transect is surveyed 12-15 times between April and October. The surveyor uses both visual encounter and presence below artificial refuges, to assess species presence/absence and to generate encounter rates. All refuges are set two weeks before the first visit to allow the animals time to begin finding and using them. Surveys are undertaken during suitable weather conditions (air temperatures between 10 and 20°C, light/no wind and no precipitation), following standard survey guidelines outlined in the National Amphibian and Reptile Recording Scheme (NARRS; see: www.narrs.org.uk).

4.1.7 Monitoring plants

4.1.7.1 Methods used to survey early successional plants in agricultural habitats

The majority of studies reviewed that collected data on early successional plants (Section 2.3.10) used either linear quadrats (e.g. 10 m length × width of ditch bank (Leng et al. 2009)) or multiple square quadrats at regular intervals along a linear transect (Jonason et al. 2011). In some studies both approaches were used, with all plants or all higher plant species surveyed in small, square quadrats and a limited number of rare plant species surveyed across whole linear plots that covered a large area (e.g. 100 × 6m; Walker et al. 2006; Walker et al. 2007). In all but two studies, the presence, cover and/or frequency of all vascular plant species were recorded; in these two the presence and cover of indicator species only were recorded. The most common response variables analysed were species richness and measures of abundance such as frequency, percentage cover, or percentage cover along a scale with defined categories such as the Braun-Blanquet scale. One study also presented data relating to plant functional traits, such as Ellenberg N which is an indication of preference for soil fertility (Hill et al. 2004). A few studies presented seed bank data that were collected from soil cores, in addition to vegetation data from quadrats (e.g. Blomqvist et al. 2003).

Marshall (2009) included important arable plant area (IAPA) scores as a response variable, derived from vegetation data collected within quadrats. The IAPA scoring system was devised
by Byfield & Wilson (2005) to assess sites for the conservation value of their arable plant assemblages, in terms of their potential to support very rare arable species with restricted distributions that are rarely recorded (Section 2.3.10). Arable species have been scored from 1-9 by (Byfield & Wilson 2005), from which cumulative species scores are calculated for each site, with different threshold values (depending on soil type) which indicate whether a site is potentially ‘important’ for arable plant conservation.

4.1.7.2 Sampling plants within focal 1km squares

Plant monitoring will take place once a year during April – September within focal 1 km survey squares. Square and linear quadrats will be allocated to AES and non-AES habitats roughly in proportion to their coverage within the 1km square, so there will be an AES vs counterfactual comparison, apart from in those squares at the ends of the AES gradients (around 0 or nearly 100% AES coverage). An algorithm to calculate proportional coverage of AES options for plant monitoring has been developed by CEH using a combination of OS Mastermap and aerial photographs for GMEP, and this will follow a similar protocol (Smart et al. 2016). All quadrat locations will be mapped using GPS coordinates and habitat landmarks in the first year of the survey, and the same fixed locations will be visited in each year of the survey.

The majority of non-linear habitats (e.g. enclosed grassland or unenclosed upland habitats) will be surveyed using up to ten nested square quadrats, with a 1 × 1 m quadrat nested in one corner of a 5 × 5 m quadrat. Vascular plant species will be identified to species and % cover estimated to the nearest 5% in the 1 × 1 m quadrat. Additional species will then be recorded in the 5 × 5 m quadrat and the % cover of each species estimated. Use of nested quadrats will allow data to be compared to the 5 × 5 m quadrats used in the NPMS scheme, some of the nests within the main ‘X’ CoS plots and the 1 × 1 m quadrats used for the HLS AES national survey (Table 11). In total up to 10 square quadrats will be surveyed per 1km square, split between types of habitat in proportion to their coverage of the square (Table 11).

Uncropped linear habitats (field margins including those under pollen and nectar and winter bird food AES options, grass buffer strips, hedgerow basal flora, linear features such as plants) will be surveyed using ten 1 × 10 m linear quadrats nested within 1 × 25 m quadrats. Vascular plant species will be identified to species and % cover estimated to the nearest 5% in the 1 × 10 m quadrat. Additional species will be recorded in the 1 × 25 m quadrat and the % cover of each species estimated. These ‘main linear’ plant transects will be placed within 200m sections of the pollinating insect and butterfly transects, where possible in the central 1m width of the 5m wide pollinator and butterfly transect.

Cropped linear habitats will be surveyed using up to ten nested linear quadrats, with the number depending on the coverage of arable habitat within the 1km square. Nested quadrats will consist of 1 × 10 m linear quadrats within 1 × 25 m quadrats within 1 × 100 m quadrats. As for the main linear transects, vascular plant species will be identified to species and % cover estimated to the nearest 5% in the 1 × 10 m quadrat, and additional species will be recorded in the 1 × 25 m quadrat and the % cover of each species estimated. Presence of additional species only will be estimated in the final 75m of the 100m long quadrat, to target rare arable species. These ‘arable linear’ plant transects will be placed in the central length of the 200m sections of the pollinating
insect and butterfly transects, where possible to coincide with the crop edge 1m of the 5m wide invertebrate transect. Main and arable linear transects may be placed next to each other, but should be separated by at least 1m between them.

<table>
<thead>
<tr>
<th>Plot</th>
<th>Size</th>
<th>Habitat coverage</th>
<th>Number per 1km square</th>
<th>Compatibility with existing monitoring schemes and datasets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enclosed or unenclosed square</td>
<td>$1 \times 1 \text{ m}$ quadrat nested corner of $5 \times 5 \text{ m}$ quadrat</td>
<td>Randomly placed not on linear features in enclosed habitats, allocated in proportion to in-field AES options</td>
<td>Up to 10 (10 enclosed and unenclosed square quadrats in total)</td>
<td>NPMS, some CoS nests of ‘X’ quadrats, some HLS survey quadrats (HLS quadrat size varies with habitat type)</td>
</tr>
<tr>
<td>Main linear</td>
<td>$1 \times 10 \text{ m}$ nested along $1 \times 25 \text{ m}$</td>
<td>Randomly placed on linear features including field margins/buffer strips, uncropped field edges (boundaries), paths and base of hedgerows, allocated in proportion to AES options that are applied to linear features</td>
<td>Up to 10</td>
<td>NPMS, CoS ‘B’ and ‘H’ plots</td>
</tr>
<tr>
<td>Arable linear</td>
<td>$1 \times 10 \text{ m}$ nested along $1 \times 25 \text{ m}$ nested along $1 \times 100 \text{ m}$</td>
<td>Randomly placed along crop edges including conservation headlands, allocated in proportion to AES arable edge options</td>
<td>Up to 10</td>
<td>NPMS; some aspects of CoS ‘H’ plots and HLS arable quadrats</td>
</tr>
<tr>
<td>Hedgerow linear</td>
<td>$1 \times 30 \text{ m}$</td>
<td>Randomly placed to sample woody component of hedgerows, allocated in proportion to length of hedgerows in square under AES management</td>
<td>Up to 5</td>
<td>CoS ‘D’ plots, HLS survey hedgerow plots</td>
</tr>
</tbody>
</table>

| Table 11 | Proposed plot sizes, shapes and habitat coverage for botanical monitoring. |

4.1.7.2 Sampling plants in surrounding eight 1km squares

In common with the invertebrate sampling, plant monitoring in the wider sampling unit will be based on a sub-sample of the eight 1km squares surrounding the focal square, using the same monitoring methods as for the focal square. Half of each of the four adjacent 1km squares will be surveyed (see Figure 26 above), with square and linear quadrats allocated to the four half squares.
at the same average density as in the focal 1km square. The use of the same intensity of sampling in the focal 1 km squares and sub-sampled adjacent squares will allow estimates of species richness to be compared, as for the invertebrates (Section 4.1.1.4).

4.2 Assessments of AES implementation

Previous monitoring of management quality and compliance to AES prescriptions has revealed wide variation in the quality of implementation and outcomes have differed considerably between and within options (Boatman et al. 2013a). Monitoring of actual resource provision or habitat quality is therefore necessary to provide a direct link between habitat provision and species responses to AES interventions.

A conceptual framework was developed for scoring the quality of habitat or resource provision arising from the implementation of options. This framework considers the attributes which could be measured, based on the literature review and previous projects that the project team have run (including Higher and Entry Level Stewardship monitoring and Campaign for the Farmed Environment monitoring), and the number and timing of visits. The framework focuses on attributes that can be determined in the field, rather than relying on farmer interviews to provide information on previous events or management.

Examples of the types of scoring are given in Appendix E, covering both long-term and newly-created habitat, in-field and boundary options, and which have value for a range of taxa:

- Nectar flower mix – e.g. of created habitat (Table E2)
- Stubble fields – e.g. of in-field habitat (Table E3)
- Conservation headlands – e.g. of boundary habitat (Table E4)
- Species-rich grassland – e.g. of long-term BAP habitat (Table E5)

Discussion of these draft scoring approaches concluded that a much simpler method of scoring was required therefore we propose the following, which may be possible to conduct during the course of other survey visits to focal survey areas:

1. To assess grasslands, margins and other flower-rich options, including nectar-flower patches and wild bird seed mixes, in terms of their benefits for pollinators & butterflies, we propose measuring just one attribute: floral resource.
2. To assess arable fields, conservation headlands, wild bird seed mixes, in term of their benefit for granivorous birds, we propose measuring two attributes: seed resource & structure.

However, if other taxa were included in the monitoring, appropriate habitat quality scoring would need to be considered. Further details of the approach taken to develop the implementation scoring examples, together with details of the attributes used in previous assessments of AES and CFE implementation (e.g. Boatman et al. 2013a; Boatman et al. 2013b), are given in Appendix E.
4.3 Estimates of number of sampling units needed including power analyses

4.4.1 The power of monitoring to detect effects of AES interventions

For a monitoring programme to have sufficient power to detect AES effects of policy interest and biological importance is one critical dimension in its design for cost-effectiveness. To that end, power analyses have been conducted, based on assumptions about expected survey results informed by real data from surveys of birds and butterflies analogous to those proposed here (Appendix F Table F2). However, partly because of the limits to the resources available and partly because appropriate “seed” data do not exist, it has not been possible to conduct power analyses for all landscape research questions of interest, across all taxa reviewed, or considering all specific potential response variables (power analyses are of counts, and may also be applicable to species richness, but do not consider diversity or percentage cover, for example). In addition, it has only been possible to consider a single analytical structure with the bird, pollinating insect and butterfly data (albeit one proven in the literature and quite likely to be applied to the data collected in a new scheme in practice): log-linear models of changes in rates of population growth with respect to management quantity within independent focal 1 km squares. Potential spill-over of mobile taxa from areas under AES intervention will be analysed using spatially-linked abundance data from adjacent areas with contrasting levels of AES intervention, so this element of AES impact would require a different model structure, which has not been tested before. Therefore, effect sizes in practice for the latter are unpredictable and conducting a meaningful power analysis in this analytical context is problematic. In addition, resources and time were only available for one set of power analyses.

The power of a scheme in practice will be determined by the variability in the data collected and the real effects of the management, as well as the influences of extraneous factors such as weather and market-driven changes in cropping, none of which is entirely predictable in advance. All of this means that the results of the power analyses described below (and in more detail in Appendix F) should be considered to be indicative, rather than definitive: power can only really be determined after data have been collected and analysed. Suggestions for levels of replication to detect spill-over effects or habitat preference in Section 4.5 are taken from previous research projects which vary in design and context compared to the proposed landscape monitoring, and are not underpinned by power analysis, so are subject to greater uncertainty.

As well as the formal, numerical power analyses, the recommendations for monitoring approaches incorporate consideration of the principles behind maximizing the potential to detect effects. Thus, (i) high replication and random sampling of sample units favours the separation of management effects from those of background conditions, (ii) maximization of the length and steepness of gradients in quantities of management among sample units favours the detection of effects of management, (iii) collecting data on population parameters that are sensitive to management effects (e.g. abundance versus species presence/absence, or species richness in addition to abundance for invertebrates) makes the rapid detection of effects more likely, (iv) consideration of secondary data that may provide insights into whether management effects have occurred, as opposed to definitive evidence (e.g. habitat selection versus change in abundance), provides additional inference both in the early stages of a monitoring programme (before any long-term effects are expected) and, ultimately, helps to interpret long-term patterns that are
found not to be statistically significant (i.e. whether this reflects low power or the lack of a biological effect). We are confident that the combination of formal and conceptual assessments of power provides the most comprehensive incorporation of the likely sensitivity of monitoring approaches that was possible within this project.

4.4.2 Estimating the effects of AES on wildlife counts: a power analysis

Estimates of power were derived using a generic power analysis approach, developed previously by CEH for research projects relating to environmental policy questions (Hails et al. 2013; Henrys et al. 2014). A full report including details of the methodology and results are given in Appendix F. In summary, the generic power approach involves simulating data relating to population change of wildlife taxa, over a range of possible numbers of sites and years, in addition to other parameters. A number of scenarios (sets of parameter values with which to generate counts) were selected by generating a random value for each parameter, from a range determined a priori to cover likely values in the proposed survey (Appendix F, Table F1). Parameter values were partly informed by data on butterflies, pollinating invertebrates and butterflies, derived from previous research projects run by CEH and BTO (Table F2). Power was simulated using Monte-Carlo simulations and tested across the range of parameter values. Generic equations were then fitted to the simulated data, which allow predictions of power for any combination of ten parameters across the scenarios, as long as the selected parameter values for any given power estimate lie within the range used to parameterise the initial simulations. The advantage of the generic approach is that any number of scenarios can be explored in relation to predicted power without the need to generate new simulations of data in each case.

The variables included in the data simulations and the generic equations are: total number of sites, total number of years, mean and variance of expected initial abundance (log scale), mean annual population trend in absence of AES management, strength of the effect of AES management, the proportion of sites with AES management, average cover of AES management at a site, a covariate, and the ratio of variance/mean (Table F1). Some of these variables (e.g. number of sites, number of years) are within the control of those designing or funding a monitoring scheme, some with vary with the species, species group or response variable chosen for assessment (e.g. mean initial abundance), and others will be properties that may vary but cannot be altered by those designing a monitoring scheme (e.g. average cover of AES management at a site). The covariate allows inclusion of a control variable, such as the cover of a particular gross habitat that might confound management effects. The variance/mean ratio relates to how over-dispersed the data are, with respect to a simple Poisson model. In order to explore the implications of over-dispersed data on power, generic equations were fitted to data simulated using negative binomial distributions, in addition to those fitted using Poisson distributions (Appendix F Figures F33-F34).
Figure 32 Power (y) as a function of the (log) expected mean abundance in year 1 (x), where counts are ‘low Poisson’. Surveys carried out at 20 (black), 50 (red), 100 (green) or 200 (blue) sites. Top left panel $\alpha = 0.01$, top centre $\alpha = 0.02$, top right $\alpha = 0.03$, bottom left $\alpha = 0.04$, bottom centre $\alpha = 0.05$, bottom right $\alpha = 0.06$. Other variables fixed at number of years = 5, variance of initial counts = 2 (log scale), mean annual trend = 0, proportion of sites treated = 0.5, average proportion of site under AES management = 0.5 (Figure F40 in Appendix F).

Power curves were generated to show the relationships with key variables (e.g. Figures 32 – 35 below). A greater range of power curves across more variables are in Appendix F (Figure F27-F43). The expected strength of the AES intervention effect strongly affects predicted power; this is shown in the multiple panels in each figure, going from low ($\alpha = 0.01$) in the top left panel to high in the bottom right panel ($\alpha = 0.06$) for each set of power curves. Alpha denotes the average difference in population trend between sites with AES intervention compared to those without, per year. The proposed monitoring uses a landscape design not previously tested, rendering it difficult to draw predictions of $\alpha$ from previous studies. For example, Brereton et al. (2007) found a 15% difference over 5 years in rates of Chalkhill blue butterfly population change at
sites with HLS management, compared to sites without AES management, using UKBMS data (i.e. per year, average $\alpha = 0.03$). However, UKBMS sites were not positioned to maximise the contrast between sites along AES gradients, as proposed here (Section 3). AES management effects may be greater under the proposed scheme, but this cannot be predicted with much certainty. Note, however, that this statistical analysis does not consider the biological constraint whereby high initial counts may reflect saturated habitat and therefore a lack of scope for population increase.

**Figure 33** Power (y) as a function of the (log) expected mean abundance in year 1 (x), where counts are ‘high Poisson’. Surveys carried out at 20 (black), 50 (red), 100 (green) or 200 (blue) sites. Top left panel $\alpha = 0.01$, top centre $\alpha = 0.02$, top right $\alpha = 0.03$, bottom left $\alpha = 0.04$, bottom centre $\alpha = 0.05$, bottom right $\alpha = 0.06$. Other variables fixed at number of years = 5, variance of initial counts = 2 (log scale), mean annual trend = 0, proportion of sites treated = 0.5, average proportion of site under AES management = 0.5 (Figure F41 in Appendix F).

Power is likely to be greater for response variables relating to high initial mean counts (Figure 33) vs. those with low initial mean counts (Figure 32). Thus, power will be greater to detect
effects of AES interventions on the abundance of more common pollinator, butterfly or bird species, and in some cases for aggregations of abundance data across multiple species (e.g. abundance of all solitary bee species). Power may not be acceptable (assuming a required power ≥ 0.8) for response variables where initial mean count is less than around 10 (examples given in Appendix F Table F2) assuming a 5 year monitoring period, even if 200 sites were set up (Figure 34).

Figure 34 Power ($y$) as a function of the number of sites surveyed ($x$), where counts are ‘low Poisson’ (X3<3, log scale). Surveys run for 6 (black), 7 (red), 8 (green) or 9 (blue) years. Top left panel $\alpha = 0.01$, top centre $\alpha = 0.02$, top right $\alpha = 0.03$, bottom left $\alpha = 0.04$, bottom centre $\alpha = 0.05$, bottom right $\alpha = 0.06$. Other variables fixed at mean initial count = 2 (log scale), variance of initial counts = 2 (log scale), mean annual trend = 0, proportion of sites treated = 0.5, average proportion of site under AES management = 0.5 (Figure F28 in Appendix F).

One option to increase power for those response variables which are expected to have low initial counts would be to monitor over a longer time period than 5 years (Figure 34).
Figure 35: Power ($\gamma$) as a function of the proportion of 200 sites treated ($\alpha$), where counts are ‘high Poisson’. Strength of coverage (B) is 0.2 (20% -black), 0.4 (red), 0.6 (green) or 0.8 (blue). Top left panel $\alpha = 0.01$, top centre $\alpha = 0.02$, top right $\alpha = 0.03$, bottom left $\alpha = 0.04$, bottom centre $\alpha = 0.05$, bottom right $\alpha = 0.06$. Other variables fixed at number of years = 5, mean initial count = 4 (log scale), variance of initial counts = 2 (log scale), mean annual trend = 0 (Figure F32 in Appendix F).

Figures 32 and 33, discussed above, are based on 50% of sites having AES interventions, analogous to a discrete AES vs. counterfactual comparison. In the proposed landscape monitoring scheme, sampling units would be allocated according to their position in either three or four classes along an AES gradient, so a quarter to a third of sites would have no or very low AES interventions. Figure 35 shows that predicted power is increased by this approach, as power increases as the proportion of AES sites goes from 0.5 to 0.75.
4.4 Estimates of the resources needed to monitor landscape-scale effects of AES

Costs were estimated for the proposed fieldwork, using the taxa protocols specified above, and the project team’s experience of running previous large, field-based monitoring schemes. Nonetheless, there is a degree of uncertainty around any resource estimates and, as discussed above (Section 4.1), for some taxa there is more uncertainty about the intensity of sampling required within a 1km square to provide estimates of abundance that are likely to be accurate, than for other taxa that have been researched more in the context of AES management. In addition, for AES implementation assessments we have provided a framework rather than detailed protocols, as deriving detailed methods across the full range of AES options that are potentially of interest is outside the scope and timescale of this project. This would need to be discussed with NE in the event of a monitoring scheme being set up. Due to this, costs around AES implementation monitoring are even less certain than those for monitoring taxa.

Costs have been estimated assuming professional field surveyors are employed. Participants at the stakeholder workshop commented that many volunteers who collect data for existing monitoring schemes are unlikely to be willing to undertake additional data collection tasks, or to collect data in intensively farmed locations where rare species are seldom recorded and a proportion of the surveying may consist of recording zero abundance (Section 2.2, Appendix B session C3 notes). This could lead to difficulties in recruiting volunteers, or a drop-off in the number of people currently volunteering. BTO experience with volunteer surveys is that interest drops off rapidly with more prescriptive and intensive sampling requirements, and that the certainty that coverage of individual survey sites can be achieved consistently can be low, even in surveys with high overall coverage and sampling effort (e.g. the BBS). The recent NPPMF project also considered the potential to use volunteers in the context of designing a new pollinator monitoring scheme (Carvell et al. 2016), and concluded that taxonomic expertise was a limiting factor in the expansion of pollinator insect monitoring by volunteers, as identification beyond broad species groups can take years to master. As discussed above (Sections 2.5 and 2.6), existing monitoring scheme data can be utilised in the context of scaling up landscape monitoring trends, but existing sites are not located to maximise the contrast along a gradient of AES gradient, existing monitoring protocols may not be intensive enough to quantify local abundance accurately, and there is a high turn-over in sites so consistent data collection cannot be guaranteed over a particular, say, 5 year period.

Annual costs of the fieldwork across the full range of taxa for which protocols were developed are presented in Table 12. These costs are modular, as the decision about which taxa to monitor is a policy one, and relates to other AES monitoring projects (e.g. the agreement-scale monitoring of the new CS scheme). The project team do not advocate that all the taxa for which costings are provided should necessarily be sampled, not least because the logistics may render such a survey unfeasible. Note also that these estimates (Table 12) are for fieldwork and directly related costs only (e.g. arranging site access, managing the surveyor teams), and do not cover the full costs of an AES landscape monitoring project (project management, data analysis, reporting and knowledge transfer are not costed). Costs included in Table 12 that relate directly to running the field survey (survey and survey team management, mileage, accommodation and data entry) have been estimated assuming that pollinating insects, butterflies, birds and plants are monitored,
in addition to AES option implementation and habitat surveying. For every additional taxon surveyed, these costs would increase as detailed in the notes for Table 12.

Where possible without compromising optimal design, the protocols were designed to survey more than one taxon on the same visit (e.g. for butterflies and bumblebees, Section 4.1) to reduce costs. However, for many taxa there is a trade-off between using optimal survey protocols, in relation to timing and the number of visits, and the capacity to combine surveys with other taxa. In addition, as more taxa are combined, this leads to a requirement for field surveyors to have a greater range of taxonomic skills, which may limit the potential pool of field surveyors. In the main, costs in Table 12 are thus modular, and do not account for possible economies of scale from combining surveys for multiple taxonomic groups. Clearly, policy decisions need to be made about survey priorities before it is possible to cost the details of integrated surveys. Further, while compromises in survey design to facilitate multi-taxa monitoring may be possible, they need to be considered carefully to maximize the quality of inference that is possible and thus to make the whole survey package genuinely cost-effective.

Table 13 shows summary (high level) estimates of costs across entire AES landscape projects, to illustrate several options that would test different research questions, depending on time-scales and the objective of the monitoring. As discussed in detail (Sections 2 and 3), we have concluded from the evidence reviewed, and from discussions with the project team, steering group and external stakeholders, that the main interest in landscape monitoring of AES interventions lies in testing whether AES effects are (1) sustained over time and (2) spill-over to affect more mobile taxa in areas of land not directly under AES management. Whether one or both of these should be monitored will be determined by policy priorities.

Table 13 shows summary (high level) estimates of costs across entire AES landscape projects, to illustrate several options that would test different research questions, depending on time-scales and the objective of the monitoring. As discussed in detail (Sections 2 and 3), we have concluded from the evidence reviewed, and from discussions with the project team, steering group and external stakeholders, that the main interest in landscape monitoring of AES interventions lies in testing whether AES effects are (1) sustained over time and (2) spill-over to affect more mobile taxa in areas of land not directly under AES management. Whether one or both of these should be monitored will be determined by policy priorities.

The landscape study design explored in Section 3 was based on a matrix of contrasting local and landscape AES gradients, with 16 or possibly 9 categories of landscape sampling units, defined on the basis of positions along the two gradients. One or two sampling units could be monitored in each position of the matrix, within an area of homogenous background landscape characteristics (an NCA). Using 20 – 50 replicate landscape sampling units to monitor spill-over would allow landscape scale monitoring to take place in two to five NCAs. However, the estimate of 20 – 50 replicates (Table 13, rows 1, 3 and 5) is quite uncertain. It was informed by Jönsson et al.’s (2015) study, in which 18 sites were selected along a gradient of landscape complexity (as opposed to a gradient of landscape AES uptake proposed here), and includes a local AES/counterfactual comparison. In their two year study, Jönsson et al. (2015) found evidence of spill-over of pollinators from wildflower strips, but they did not survey other taxa, or monitor change in populations over time. Thus, the best available comparable studies have been used to provide an indication of likely replication for those project options that were not informed by power analysis, but there are substantive differences between existing studies and the proposed landscape survey design. The replication estimate of 100 – 200 1km squares for monitoring focal squares only (Table 13 option 2) is derived from the generic power analysis (Section 4.4.2 above) based on number of sites across a national monitoring scheme, as data were not available to parametise models for a power analysis based on NCA groupings. Section 3 demonstrated only one replicate of some cells of a four × four matrix of contrasting local and landscape gradients would probably be available within each NCA. If this matrix design were
chosen, the 100 – 200 replicates would probably thus need to be distributed across six to twelve NCAs (for a four × four matrix), for practical reasons.

This project was commissioned in the context of a 5 year monitoring timescale for the new CS AES (Table 13a), but as funding priorities may change shorter timescales have also been costed (Table 13b), while recognising that evidence of population change is unlikely to be detected in less than 5 years. Costs were estimated to include monitoring pollinating insects, butterflies, birds and plants, and also AES option implementation and habitat mapping, in addition to project management, reporting and a limited about of knowledge transfer (Table 13). The modular annual costs in Table 12 can be used to compare the cost of adding specific additional taxa to the monitoring project possibilities in Table 13 (while taking into consideration that costs of managing the field team, mileage etc. will also increase with additional taxa surveyed). The Glastir Monitoring and Evaluation Programme (GMEP) is the only other UK AES monitoring project to survey multiple taxa (including plants, butterflies and birds), in addition to other environmental variables (e.g. soil carbon content, greenhouse gases and landscape visual quality). For comparison, GMEP costs between £1.3 and £1.6 million a year, to survey on average 75 1km squares a year.
<table>
<thead>
<tr>
<th>Task / taxa monitored</th>
<th>Survey methods / notes</th>
<th>20 sites</th>
<th>50 sites</th>
<th>100 sites</th>
<th>200 sites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Focal 1km only</td>
<td>Spill over: 3 x 3 km</td>
<td>Focal 1km only</td>
<td>Spill over: 3 x 3 km</td>
</tr>
<tr>
<td>Mapping habitats and options in 1km squares, year 1 only</td>
<td>Survey onto tablets using Arc (tablets also used for other field data collection)</td>
<td>£38,000 £96,000</td>
<td>£68,000 £186,000</td>
<td>£136,000 £372,000</td>
<td>£236,000 £672,000</td>
</tr>
<tr>
<td>Pollinators (bumblebees, solitary bees, hoverflies) and butterfly fieldwork</td>
<td>2 × 1km transects including floral resource indices, set 6 pan traps, collect 24 hours later. 4 times May - Aug</td>
<td>£40,150 £100,450</td>
<td>£100,300 £250,900</td>
<td>£200,600 £501,800</td>
<td>£401,200 £1,003,600</td>
</tr>
<tr>
<td>Pollinators laboratory work</td>
<td>Identification of solitary bees, bumblebees and hoverflies from pan traps</td>
<td>£52,800 £158,400</td>
<td>£132,000 £396,000</td>
<td>£264,000 £792,000</td>
<td>£528,000 £1,584,000</td>
</tr>
<tr>
<td>Moths</td>
<td>Set and collect 6 light traps per 1km, identify macro-moths in field. Twice during June &amp; July</td>
<td>£26,250 £78,750</td>
<td>£65,000 £195,000</td>
<td>£130,000 £393,000</td>
<td>£260,000 £1,650,000</td>
</tr>
<tr>
<td>Bird breeding season &amp; terrestrial mammals monitored under BBS</td>
<td>3 × 1km transects, 4 times April - July</td>
<td>£20,000 £60,000</td>
<td>£50,000 £150,000</td>
<td>£100,000 £300,000</td>
<td>£200,000 £600,000</td>
</tr>
<tr>
<td></td>
<td>3 × 1km transects, 4 times November - March</td>
<td>£20,000 £60,000</td>
<td>£50,000 £150,000</td>
<td>£100,000 £300,000</td>
<td>£200,000 £600,000</td>
</tr>
<tr>
<td>Mammals - field vole and harvest mouse</td>
<td>4 × 100 m field sign transects, one visit for field voles in autumn and harvest mice in December</td>
<td>£10,000 £40,000</td>
<td>£25,000 £100,000</td>
<td>£50,000 £200,000</td>
<td>£100,000 £400,000</td>
</tr>
<tr>
<td>Mammals - dormouse</td>
<td>Artificial nest boxes set and checked twice per year</td>
<td>£15,138 £60,414</td>
<td>£37,845 £151,035</td>
<td>£75,690 £302,070</td>
<td>£151,380 £604,140</td>
</tr>
<tr>
<td>Mammals - hedgehog</td>
<td>Baited footprint tunnels set and checked over 4 consecutive days</td>
<td>£25,160 £75,021</td>
<td>£62,820 £187,541</td>
<td>£125,640 £375,083</td>
<td>£251,280 £750,166</td>
</tr>
<tr>
<td>Bats</td>
<td>Setting and collecting static bat detectors. Repeat each month May to Aug</td>
<td>£23,914 £35,875</td>
<td>£61,117 £91,688</td>
<td>£122,900 £184,375</td>
<td>£245,800 £368,750</td>
</tr>
<tr>
<td>Bats</td>
<td>Analysis of acoustic signals</td>
<td>£12,000 £24,000</td>
<td>£30,000 £60,000</td>
<td>£60,000 £120,000</td>
<td>£120,000 £240,000</td>
</tr>
<tr>
<td>Reptiles (adults)</td>
<td>2 × 1km transects and 30 refugia sampling</td>
<td>£45,280 £88,000</td>
<td>£113,200 £220,000</td>
<td>£226,400 £440,000</td>
<td>£452,800 £880,000</td>
</tr>
<tr>
<td>Amphibians</td>
<td>6 × pond surveys, refugia along 100m transects</td>
<td>£30,000 £45,000</td>
<td>£75,000 £112,500</td>
<td>£150,000 £225,000</td>
<td>£300,000 £450,000</td>
</tr>
<tr>
<td>Plants</td>
<td>Square and linear quadrats, up to 30 per 1km square depending on habitats. Once a year</td>
<td>£55,500 £166,000</td>
<td>£138,000 £413,500</td>
<td>£276,000 £872,000</td>
<td>£552,000 £1,654,000</td>
</tr>
</tbody>
</table>
## Estimated annual fieldwork costs

<table>
<thead>
<tr>
<th>Task / taxa monitored</th>
<th>Survey methods / notes</th>
<th>20 sites</th>
<th>50 sites</th>
<th>100 sites</th>
<th>200 sites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Focal 1km only</td>
<td>Spill over: 3x3km</td>
<td>Focal 1km only</td>
<td>Spill over: 3x3km</td>
</tr>
<tr>
<td>Survey and survey team management</td>
<td>Includes arranging site access, managing surveyors and survey schedule, data quality</td>
<td>£57,250</td>
<td>£80,750</td>
<td>£57,250</td>
<td>£80,750</td>
</tr>
<tr>
<td>AES option implementation scoring</td>
<td>Rapid scoring assessment, done at same time as botanical survey</td>
<td>£30,500</td>
<td>£121,000</td>
<td>£75,500</td>
<td>£301,000</td>
</tr>
<tr>
<td>T&amp;S, mileage, accommodation</td>
<td></td>
<td>£27,467</td>
<td>£82,400</td>
<td>£68,667</td>
<td>£206,000</td>
</tr>
<tr>
<td>Data input for taxa not recorded on tablets in field</td>
<td>Birds, bats, mammals, herptiles, potentially pollinators and butterflies</td>
<td>£7,500</td>
<td>£22,500</td>
<td>£18,750</td>
<td>£56,250</td>
</tr>
</tbody>
</table>

### Notes:

1. Mapping habitats and options. Assumes AES options mapped in focal 1km squares only, habitats mapped in focal and adjacent squares. Equipment costs include ruggedised tablets, also used for surveying plants and potentially other taxa. Mapping and tablet costs in year 1 only.
2. Pollinators and butterfly fieldwork. Equipment costs include buying pan trap stations, equipment costs in year 1 only.
3. Pollinator laboratory work. Identification of solitary bees, bumblebees and hoverflies to species from pan trap samples. Equipment costs include sample tubes and alcohol for sample storage.
4. Moths. Identification of macro-moths to species in field. Equipment costs include heath light traps, equipment costs in year 1 only.
5. Mammals. Equipment costs include nest boxes, footprint tunnels and bait, all costs but bait in year 1 only.
6. Bats. Equipment costs include static bat detectors, equipment costs in year 1 only.
7. Reptiles. Equipment costs include refugia, equipment costs in year 1 only.
8. Plants, surveyed once per year during April - September. Data recorded on ruggedized laptops (laptops costed in habitat survey as plant surveyed at same time in year 1). Minor equipment costs for quadrats and drop discs.
9. Survey team management. NB Does not include project management, data analysis, reporting and knowledge transfer. Assuming pollinators and butterflies, birds and plants surveyed. Would increase 5% per extra taxa surveyed. Equipment costs included here are basic fieldwork equipment (clipboard, PPE, limited number of cameras and GPS units).
10. AES option implementation assessment. Note framework of methods for this (Section 4.2) are indicative and for a few options only; costs will vary depending on final methodology.
11. T&S, mileage, accommodation. Costed assuming plants, birds, pollinators and butterflies surveyed. Additional 20% to be added for each extra taxa surveyed.
12. Data input. Costed assuming data input for birds, pollinators and butterflies (habitat and plant data entered directly into ruggedized tablets in field). Data entry costs would increase by 20% for every additional taxa surveyed.

**Table 12** Estimated annual fieldwork costs, per taxonomic group with a specific sampling method and associated key tasks (e.g. assessing AES option implementation and arranging access to landscape sampling units), using the protocols in Section 4.1. These costs relate to fieldwork and directly associated costs only. Estimates for project management, data analysis, reporting and knowledge transfer are not included (see Table 13). Sites refer to individual 1km squares, with (spill-over: 3×3km) or without (focal 1km only) a surrounding buffer of 8 further km squares forming a block of 3×3km.
Table 13 Summary of full project costs for monitoring AES landscape-scale effects, the research question addressed and notes on pros and cons of each possibility for (a) a 5 year monitoring scheme, the timescale of the new CS AES and (b) monitoring over shorter time-scales. Costs assume monitoring of pollinating insects, butterflies, birds and plants, in addition to AES implementation and habitat mapping. Sites refer to individual 1km squares, with or without a surrounding buffer of 8 further km squares forming a landscape sampling unit of 3×3km.

<table>
<thead>
<tr>
<th>Description</th>
<th>Research question addressed</th>
<th>Estimated replication (approx. 9 sampling units per NCA)</th>
<th>Replication estimate informed by</th>
<th>Number of years</th>
<th>Cost estimate (£000s)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Monitoring pollinators, butterflies, birds, plants, habitat and AES option implementation in focal 1 km squares and adjacent squares that make up 3 × 3 km landscape units, across a matrix of contrasting 1 × 1 km and 3 × 3 km AES gradients, for five years using protocols described in Section 4.2</td>
<td>Do AES interventions lead to an increase (or reduced decrease) in abundance of key mobile taxa that spills over to alter abundance on non-AES land within landscape sampling units (3 × 3 km), and are any changes sustained over time?</td>
<td>20 - 50 landscape sampling units (3 × 3 km)</td>
<td>Previous research study</td>
<td>5</td>
<td>£4,590 - £9,550</td>
<td>High uncertainty in replicate estimate as spill-over previously shown in single year for pollinators only in two studies; not assessed as spill-over in terms of temporal change, or for other taxa</td>
</tr>
<tr>
<td>2. Monitoring pollinators, butterflies, birds, plants, habitat and AES option implementation in focal 1 km squares only across a matrix of contrasting 1 × 1 km and 3 × 3 km AES gradients, for five years using protocols described in Section 4.2</td>
<td>Is the abundance of key mobile taxa within 1 km square local sampling units affected by AES interventions at local and landscape (3 × 3 km) scales, and are any effects sustained over time?</td>
<td>100 - 200 1km squares</td>
<td>Power analysis</td>
<td>5</td>
<td>£6,705 - £12,810</td>
<td>1) Replication required depends on taxa and response variable, see power analysis Section 4.3. 2) Monitoring only in focal 1km squares, no assessment of local and landscape AES effects on spatially-linked estimates of abundance.</td>
</tr>
<tr>
<td>Description</td>
<td>Research question addressed</td>
<td>Estimated replication</td>
<td>Replication estimate informed by</td>
<td>Number of years</td>
<td>Cost estimate (£000s)</td>
<td>Notes</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>---------------------------------</td>
<td>----------------</td>
<td>----------------------</td>
<td>-------</td>
</tr>
<tr>
<td>3. Monitoring pollinators, butterflies, birds, plants, habitat and AES option implementation in focal 1km squares and adjacent squares that make up 3×3km landscape units, across a matrix of contrasting 1×1km and 3×3km AES gradients, for a single year</td>
<td>Do AES interventions alter habitat selection by key mobile taxa at local and landscape scales?</td>
<td>20 - 50 landscape sampling units (3×3km)</td>
<td>Previous research studies</td>
<td>1</td>
<td>£996 - £2,060</td>
<td>1) Risk of not collecting robust data due to low abundance caused by atypical weather in single year of project, especially for invertebrates and birds. 2) No information about whether any effects seen in single year will be sustained over time, may just assess temporary habitat selection. 3) Uncertainty around replication estimates as previous study designs differ from proposed methodology.</td>
</tr>
<tr>
<td>4. Monitoring pollinators, butterflies, birds, plants, habitat and AES option implementation in focal 1km squares only, across a matrix of contrasting 1×1 km and 3×3 km AES gradients, for a single year</td>
<td>Is the habitat selection of key mobile taxa within 1 km square local sampling units affected by AES interventions at local and landscape (3×3 km) scales?</td>
<td>100 - 200 1km squares</td>
<td>Previous research studies</td>
<td>1</td>
<td>£1,450 - £2,752</td>
<td>1) Risk of not collecting robust data due to low abundance caused by atypical weather in single year of project. 2) No information about whether effects seen in single year will be sustained, may just assess temporary habitat selection. 3) Uncertainty around replication estimates as previous study designs differ from methodology developed here.</td>
</tr>
<tr>
<td>5. Monitoring pollinators, butterflies, birds, plants, habitat and AES option implementation in focal 1km squares and adjacent squares that make up 3×3km landscape units, across a matrix of contrasting 1×1km and 3×3km AES gradients, for 3 years</td>
<td>Do AES interventions lead to an increase (or reduced decrease) in abundance of key mobile taxa, which spills over to alter abundance in adjacent areas, and are changes in abundance sustained for &gt; 1 year?</td>
<td>20 - 50 landscape sampling units (3×3km)</td>
<td>Previous research studies</td>
<td>3</td>
<td>£2,794 - £5,803</td>
<td>1) Risk of not collecting robust data in single atypical year reduced by three years of fieldwork. 2) Unlikely to determine whether changes in rate of population increase/decline occurring in just three years, but will provide evidence about whether short-term changes in abundance are sustained over more than a year. 3) Uncertainty around replication estimates as previous study designs differ from methodology developed here.</td>
</tr>
<tr>
<td>6. Monitoring pollinators, butterflies, birds, plants, habitat and AES option implementation in focal 1km squares only, across a matrix of contrasting 1×1km and 3×3km AES gradients, for 3 years</td>
<td>Is the abundance of key mobile taxa within 1 km square local sampling units affected by AES interventions at local and landscape (3×3 km) scales, and are effects sustained for &gt; 1 year?</td>
<td>100 - 200 1km squares</td>
<td>Previous research studies and power analysis</td>
<td>3</td>
<td>£4,077 - £7,781</td>
<td>1) Risk of not collecting robust data in single atypical year reduced by three years of fieldwork. 2) Unlikely to determine whether changes in rate of population increase/decline occurring in just three years, but will provide evidence about whether short-term changes in abundance are sustained over more than a year. 3) Uncertainty around replication estimates as previous study designs differ from methodology developed here, and estimates differ with response variable and taxa monitored. 4) Monitoring only in focal 1km squares, no assessment of local and landscape AES effects on spatially-linked estimates of abundance.</td>
</tr>
</tbody>
</table>
4.5 Species monitoring strategy: conclusions

Detailed protocols were designed for several species groups, based on surveying sampling units with contrasting local and landscape AES gradients (Section 3). Protocols are provided for pollinating insects, butterflies, birds, bats and plants (identified as candidates for landscape-scale monitoring of AES effects by NE) and also moths, terrestrial mammals, reptiles and amphibians (included as a result of the workshop scoring exercises, Section 2.2). The extent to which monitoring has been undertaken previously in the context of AES differs between taxa, so knowledge of the sensitivity and power associated with possible survey approaches for birds and butterflies is far greater than that for amphibians, for example. Decisions about which species groups to monitor will depend on policy priorities, and the extent to which taxa are surveyed in other AES monitoring projects. A framework to monitor AES implementation was also developed, informed by the project team’s previous experience of scoring ELS implementation.

A generic power analysis approach was used, based on modelling population change of mobile organisms in 1km squares that contained varying proportions of AES intervention, which was parameterised using data from previous monitoring of birds, butterflies and pollinating insects. Power was greater for those response variables that are likely to have an initial mean count larger than around 10, so for some taxa it is unlikely that the changes in the abundance of less common species will be detectable within a 5 year monitoring period, and the use of aggregated response variables may be necessary. The likely replication needed will thus depend on policy decisions regarding which taxa to monitor, and how long the monitoring is likely to continue. The power graphs (Section 4.3 and Appendix F) can be used to provide estimates of the likely replication needed, once these policy decisions have been made. The power analysis addressed one potential monitoring scenario (changes in populations in 1km focal squares). Estimates of replication needed for other scenarios (e.g. spill-over between adjacent squares) have been informed only by the limited available previous studies, so are subject to greater uncertainty.

Indicative estimates of the resources needed to monitor each taxon in the field are provided (Section 4.4), based on the protocols (Section 4.1). These taxon fieldwork costs are modular to enable comparison between taxa, and do not include the full cost of running a monitoring project. Full costs are estimated for six monitoring scheme scenarios, based on monitoring pollinating insects, butterflies, birds and plants to assess (a) spill-over between adjacent squares vs. (b) population change in focal squares only, over various timescales, including the 5-year monitoring suggested for this scoping study.
5 Review of analytical techniques for landscape-scale biodiversity data

5.1 Summary of strengths & weaknesses of modelling methods for landscape-scale detection of AES effects

The design of any monitoring programme should include a strategy for analysis of any data collected. A range of statistical methodologies that could be applied to data collected following the proposals in Section 4 are discussed below, with reference to previous use in AES attribution studies. It is important to note that the methods discussed are not discrete or alternative approaches, and it is likely that a combination of methods will be used in any final analytical strategy. Data collected on different taxonomic groups will have different properties and may require different analytical approaches.

The first part of this chapter draws together the strengths, weaknesses and opportunities for applying each class of method to the objective of detecting and then predicting landscape-scale effects of AES. These conclusions build on the more comprehensive review of each method that forms the rest of the chapter.

5.1.1 Stratified random sampling

Given the emphasis of the project on testing for landscape-scale effects of AES uptake on particular taxon groups, recommendations for field survey emphasise a robust design where gradients of focal and landscape-scale uptake cover as great a range of AES intervention as possible, and are crossed to ensure independence of their effects. The ability to upscale then critically depends on the precision and transferability of the model derived from analysis of the crossed design. This in turn relies on the sample being sufficiently representative of the national range of habitats and landscapes and how their attributes condition response to AES uptake. A stratified random sample of 1km squares would not be guaranteed to provide the necessary contrast between low and high AES uptake at each scale and so would be a weaker design for hypothesis testing. However a stratified random sample would allow for robust inference and upscaling of the average species response in areas that could be stratified partially by AES uptake. There are high risks with both strategies but the analytical challenge of testing explicitly for landscape-scale effects and using a derived model to upscale is best served by a carefully designed sample with attendant field survey.

Ability to detect or predict landscape-scale effects: Potentially high but not as high as a designed, crossed gradient study of the kind we recommend.

5.1.2 Bayesian Belief Networks (BBN)

These methods are supported by a large literature and are well developed. They have the advantage that a model can be built using only existing knowledge and data with no recourse to new field measurements. However, therein lies their weakness. In BBN the effect sizes that estimate the relationship between say AES uptake and species response are negotiable probabilities derived from prior existing knowledge about the likely effect of the option conditioned by all the many other factors such as quality of implementation and landscape variables. This means that BBN on their own cannot directly test the hypothesis that the new and existing AES in England generate landscape-scale impacts on wildlife, and
fundamentally do not represent monitoring of wildlife responses. At some point the predictions from the BBN would need to be confronted by real data. New observations of species abundance change then play a critical role in refining the BBN. Data-gathering could be therefore be targeted at areas where predictions of impacts on species’ abundance took on a particular value; low, high, moderate. The amount of test data could be much less than a full survey across the designed matrix but the approach carries very high risk since if the BBN is substantively wrong then a set of observations may simply prove the BBN model to be flawed yet with no explanatory variables having been recorded, little understanding could be gained to add to the evidence of AES effects.

Compared to analysis of observations based on a robust, crossed design, it is clear that a BBN coupled with observations of species’ abundance for model testing would be a weaker approach. Relying solely on a BBN with no observational validation data would in no way provide the evidence required to measure and understand the landscape-scale impacts of AES in England.

**Ability to detect or predict landscape-scale effects**: Low.

### 5.1.3 Habitat Suitability modelling and use of existing models

A substantial literature exists illustrating the development and application of habitat suitability models. A highly desirable feature of such models is that they should be able to predict changes over time, yet many empirical models focus on representing spatial patterns and assume that these translate into changes in time. Moreover, a significant shortcoming in light of the goals of this project is that habitat suitability models do not explicitly include AES uptake and quality of implementation as explanatory variables and even if they did it would require a degree of belief in lieu of evidence that the models were able to reproduce the dynamic landscape-scale responses of target biota among focal English NCAs. The issue of landscape-scale benefits is especially critical. The focus is on landscape-scale effects and so models need to include the additive or multiplicative effect of AES options on species response as a function of differences in density of uptake across a larger network of habitat patches and grid squares. Where they exist for relevant taxa, the principal use of habitat suitability models should be in providing guidance on critical explanatory variables to measure in the field or to derive from existing map and EO (earth observation) products.

**Ability to detect or predict landscape-scale effects**: Low but likely to vary between taxa.

### 5.1.4 Remote sensing

Remotely sensed products offer potentially large benefits as sources of covariates for modelling landscape-scale effects of AES options on species’ abundance. Many products are freely available or are likely to be available for no or low cost in the near future. Also newer satellite coverage is at finer resolution than older coverage. However, newer technologies are often untested and costly. Moreover, newer products that, for example, estimate Net Primary Production and soil moisture are the result of complex post-processing such that they are also modelled variables with attendant uncertainties that will inevitably reduce the predictive power of an EO-derived variable.
Satellite-derived products are useful for deriving estimates of habitat diversity and composition but often have insufficient resolution to detect the fine-grained structural changes in habitat conditions that are important for target taxa and that result from AES option implementation. Remotely sensed map products also come with known classification errors. These could be included in a model of the effects of AES as informative priors. In effect this would introduce measurement error onto the covariate of remotely sensed habitat type. This would propagate uncertainty onto the predicted species response reflecting uncertainty in the identification of the habitat in any one place (see Lunn et al. 2013 for a description of measurement error modelling in a Bayesian context). Indeed, more relevant prior uncertainties could be derived by comparing habitat composition and extent mapped in the focal and landscape sample and comparing these field maps with Land Cover Map (LCM) coverage to generate survey-specific misclassification rates for the habitats in the sample. This uncertainty could be explicitly factored into the use of LCM as a source of habitat map variables when up-scaling an AES effects model to the national scale. The important message here is that greater uncertainty on the model input data will generate higher uncertainty on model predictions.

Notwithstanding their weaknesses, products such as LCM, Sentinel imagery and LCM Crop map have a key role to play in providing contextual information on habitat composition and structure. Other information sources such as Lidar provide the potential to measure the width and density of linear wooded features.

**Ability to detect and predict landscape-scale effects:** Will vary with focal taxon, product and attribute required. Cheaper than field survey and high potential for cost-effective application but explanatory power will generally be lower than field survey (Rhodes et al. 2015).

### 5.1.5 Occupancy Detection models

These methods are relevant to the analysis of existing national recording scheme data where there is known to be appreciable spatial and temporal variation in recording effort. They are increasingly applied within a Bayesian framework. A model of the data collection process is typically embedded alongside models of the fixed effects of interest and the random effects that account for sample structure. Modelling the data collection process is an essential requirement if variation in recording effort is not to be confounded with other drivers of changing abundance, which would include AES option impacts. These methods have seen rapid recent development and application to opportunistic biological records (Isaac et al. 2014; Woodcock et al. 2016) and also now underpin a number of wildlife trend indicators for the UK ([http://jncc.defra.gov.uk/page-6850](http://jncc.defra.gov.uk/page-6850)).

These methods are required where analysis of opportunistic recording data is undertaken, for example in 1km squares that coincide with the designed sample or in the wider countryside where trends in these data can be compared with model predictions of AES impact. Analysis of opportunistic data to detect AES impacts can be attempted particularly for groups of species not already well covered by systematic citizen science survey schemes. However, the extent of the spatial and temporal biases in these data will need to be assessed tactically since occupancy detection methods do not offer a universally reliable correction for such biases; if they did then logically there would be no need for data. Initial sub-setting may therefore be needed to remove regions typified by the most extreme biases in recording effort (Woodcock et al. 2016). The general principle is carefully to consider whether the ‘missing data mechanism’ can be ignored (see Gelman & Hill 2007; Lunn et al. 2013 for discussion).
is, Bayesian imputation methods can be applied in most circumstances but if variation in recording effort leads to gaps in recording that are strongly confounded with covariates such as management impacts, climate, geography and time then imputed values will be increasingly unrepresentative and uncertain the more severe the non-random nature of the missing data.

**Ability to detect and predict landscape-scale effects:** Potentially high for certain, limited taxa for which structured monitoring is not available. Some proven examples already exist. Where unquantified spatial and temporal variation in recorder effort is present these techniques are essential for effective signal attribution.

### 5.1.6 Generalised Linear Models and Mixed Models (GLM and GLMM)

Many published examples attest to the popularity of these methods for quantifying the relationship between a response variable and a series of fixed effects where the response variable and predictors are organised hierarchically (Gelman & Hill 2007). GLMM have also been widely applied in the detection of AES impacts including landscape-scale effects and legacy effects of previous scheme uptake. As such, GLMM offer an obvious and proven framework for model building and model application.

In light of the recommended sampling design that brings together crossed gradients of focal and landscape-scale uptake, we consider how GLMM might be formulated to model landscape-scale responses to AES impacts and quality of implementation starting from the simplest model focussing on field or 1km square-scale uptake and impacts only, to a model that includes the effect of varying density and quality of uptake in the surrounding landscape on the focal 1km square (Table 5).

No matter how sophisticated the modelling approach, a number of issues pertain that will reduce detection and modelling of the full effect of AES options on species’ responses. These are as follows:

1. The review of evidence for AES effects on pollinators highlighted the finding that landscapes of intermediate complexity were most responsive to intervention. However, this is probably a simplification and the conditioning effect of landscape may well vary with taxon group and associated options, as well as with the absolute levels of minimum and maximum complexity in a given region. Since we know these effects are plausible, we can attempt to test relevant hypotheses while also managing expectations for medium-term AES impacts in the most and least complex landscapes.
2. If rare species are associated with few recorded presences, then analysis of these data are more likely to return non-significant effects.
3. While plausible, there is very little evidence for the reinforcing effects of multiple options on species’ abundance. Therefore, it is possible that these synergistic impacts will be missed unless hypotheses are constructed that carefully combine uptake of groups of favourable options. However, any design considering the combined effects of multiple options specifically may entail different AES gradients to those selected to maximize contrast in the provision of individual resource types. Hence, a single monitoring design might not be able to deliver cost-effective evaluation of both simple AES effects and synergistic ones.
4. Population changes may not be explained well by AES uptake, quality of implementation and landscape factors, because other factors are more important but are not measured. This could, for example, include predation on ground-nesting birds.
5. While it should be possible to construct a sampling design based on crossing, replication and randomisation of AES uptake in focal and landscape 1km squares, there is no guarantee that other controlling variables can be equivalently crossed and therefore made independent. Thus, while this would not necessarily be a focus for any AES monitoring programme, it should be noted that these variables can be incorporated as controls, but not necessarily tested in respect of their impacts on focal taxa.

6. If there are many control or background influences of biological interest, the monitoring sample size may be too small compared to the number of effects potentially to be estimated, so it is unlikely that parameters for all these effects will be estimable in a single model. Hence, it is recommended that non-AES variables are considered only as controls, and that formal tests of their effects are left to other studies.

**Ability to detect and predict landscape-scale effects:** Proven to be high for detection of AES effects in some groups. Prediction is also possible but the transferability of the resulting model will depend on how well represented gradients of other controlling factors were in the model-building sample.

**5.1.7 Structural Equation Models (SEM)**

SEM have rapidly increased in popularity in the ecological literature (Grace 2006; Shipley 2009) and software for their implementation is now widely available. SEM can also be readily executed in a Bayesian framework (Clough 2012; Smart *et al.* 2014). The appeal of SEM is they can combine all the benefits of a mixed modelling (GLMM) approach with estimation of direct and indirect effects. That is they test the hypothesis that a response variable is conditionally independent of another variable based on the status of a mediating variable. Such relationships are common in ecology (Figure 36).

*Figure 36* a) A standard GLM or multiple regression approach where explanatory variables are treated as statistically independent and have detectable effects on the response variable, b) the same data set seen through the SEM lens with an additional hypothesis that habitat condition change is explained by AES uptake ($\beta_3$).

In this example, it is inevitable that AES uptake and habitat condition change will not be independent of one another. Analysing them as such in a standard GLM would risk underestimating the effect of each variable and more importantly not deriving a model faithful to the probable mechanism involved whereby AES uptake leads to habitat condition change which in turn provides resources and conditions favourable to the focal species. A recent example was described by Woodcock *et al.* (2016) who explored the expected collinearity between area of Oil Seed Rape and application of neonicotinoid pesticides when analysing changes in bee abundance across Britain.
In the SEM formulation (Figure 36b) the direct effect of AES uptake on species response can be estimated alongside the direct effect of uptake on habitat condition and the direct effect of habitat condition on species response. This allows estimation of the indirect effect of AES uptake on species response via the mediating variable of habitat condition change. When covariates have normally distributed errors and can therefore be centred and standardized to zero mean and 1 standard deviation then the indirect and total effects can be simply calculated by multiplication of the standardized regression coefficients which are also interpreted as effect sizes ($\beta_1$ to 3 in Figure 36b). Thus the indirect effect of AES uptake on species response is $\beta_3 \times \beta_2$ and the total effect of AES uptake is $\beta_1 + (\beta_3 \times \beta_2)$.

Using SEM, the shared variation between AES uptake and habitat condition change is subject to a hypothesis test derived from a mechanistic explanation about their collinearity rather than this being a nuisance attribute of the dataset. Where possible causal/correlative relationships of this kind can be associated with the variables in the GLMM draft model schema (Table 14) an SEM framework can be used to test for them (Smart et al 2014; Grace et al 2016).

**Ability to detect and predict landscape-scale effects:** Potentially high. When developed in a Bayesian framework, prediction is also straightforward (Gelman & Hill 2007; Ward et al. 2016).

### 5.2 Conclusions from summary of strengths and weaknesses of analytical techniques

The project objectives require that analytical options are scoped with a view to detecting the impact of AES options at the landscape scale. This involves producing the best evidence for the effects of AES options and then up-scaling these relationships to the wider countryside. Thus, analytical options coalesce around three linked tasks: hypothesis testing, model building and model application.

Quantifying the relationship between AES uptake, quality of implementation and species abundance can be written as a testable hypothesis. The strongest test will be based on a designed sampling scheme where gradients of AES option uptake at a focal and a surrounding landscape scale are made independent and where each gradient varies from highest to lowest uptake (Section 3). Species abundance is then measured along these crossed gradients and change in abundance modelled in terms of the main and interacting effects of focal and landscape-scale uptake. The test of the landscape-scale effect of AES uptake and quality of implementation on measured species response is based on inspecting the size and sign of the model parameters. This model can be applied in predictive mode to new areas, subject to its fit being sufficiently good. Knowing the values of the inputs to the model – AES uptake, quality of implementation and habitat condition change - allows prediction in these new areas hence the model can be used to upscale. Note that whilst such a model can be derived and applied for prediction, if the designed sample used for hypothesis testing and model building is substantially different in character from other parts of England, the model may have low transferability and therefore perform poorly.

The most useful model for up-scaling will be the one with the least uncertainty and maximum transferability to new areas. This will result from a strong hypothesis test dependent on a robust sampling design, but also on a sampling domain that is sufficiently representative of the wider landscape. So model performance is likely to be improved by the inclusion of other
covariates known to influence biological responses to changing management (Section 2.4 Table 5b). These factors should therefore also feature in the hypothesis testing and model building phase, but they are numerous. Testing their main and interacting effects alongside AES uptake in focal and landscape squares may not be supported by the sample size being considered. Therefore, a careful process of model and covariate selection will be required, especially if a small (20-50) number of sites are surveyed. A number of these wider landscape factors are likely to co-vary with option uptake. For example, some AES options are specific to, or much more likely to be taken up in, upland landscapes where there is also more semi-natural habitat while linear features other than watercourses are less common. These factors should be controlled as far as possible in the sampling design. Bayesian modelling methods have some merit under such circumstances because they explicitly quantify the uncertainty attaching to all effects and are thought to be robust to a degree of collinearity (Woodcock et al. 2016) and small sample sizes (Lee & Song 2004). However, high collinearity will inevitably mean that shared variance cannot be uniquely attributed to either of the correlated variables because the analysis wrongly assumes them to be independent.

Where nuisance correlation between variables can be interpreted as a plausible causal relationship, such that one factor co-varies with another through a driver-response relationship, then SEM techniques provide a way of estimating such causal linkages. For example, where AES uptake is hypothesised to drive a species’ response via changes in ecological conditions then SEM can estimate the indirect effect of AES uptake on species response via its correlation with the quality of option implementation. The ability to interpret the results of the SEM as causal depend upon having a robust crossed, replicated and randomised sampling design rather being attributable to use of the SEM technique itself.

Effort should be directed toward building the best model of landscape-scale AES effects. This requires a robust design (Section 3), control of other conditioning factors as far as possible and measurements of species abundance and change in ecological conditions driven by option implementation (Section 2.4, Table 5a). This suggests that subsequent application of the resulting model across the wider landscape will require measurement of the same range of field measurements as model inputs. The modelling approach should, however, be flexible enough to allow a number of models to be constructed in the hypothesis testing phase. More costly field survey measurements could be deliberately excluded as model covariates in favour of less costly proxies derived from existing map sources and EO datasets (e.g. Section 2.4, Table 5b). A comparison of the performance of the models with field-surveyed covariates versus models with only map and EO-derived covariates would then quantify the trade-off between reduced cost of covariate measurement versus reduction in model precision when used for prediction of AES impacts outside of the model-building domain and at the national scale. This systematic approach to the assessment of the role of map-based and EO products will ensure that their use is maximised, but also that their capacity to convey the levels of ecological detail present in more costly field measurements is explicitly tested.

The increasing popularity and sophistication of Bayesian methods means that GLMM, SEM and occupancy-detection models can now all be combined into a single modelling framework (Maxwell et al. 2016). The additional benefits of using Bayesian modelling include the ease with which predictions can be made as well as in transparently conveying the uncertainties attaching to all parameters. These methods coupled with a statistically robust sampling scheme offer the best opportunity for detecting landscape-scale effects of AES uptake on species abundance and producing up-scalable models that have least uncertainty.
<table>
<thead>
<tr>
<th>Model schema</th>
<th>Model description in words</th>
<th>Relevance to landscape-scale effects</th>
<th>Strength of evidence for landscape-scale effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Sr_f = \beta_1 * U_f$</td>
<td>Species response is a function of %AES uptake in a 1km square or between in and out of option patches in a 1km square ($U_f$).</td>
<td>The model does not combine the effects of uptake in a focal square with uptake in surrounding squares and so is not relevant to landscape effects.</td>
<td>Does not answer question about landscape effects.</td>
</tr>
<tr>
<td>$Sr_f = (\beta_1 * U_f) + (\beta_2 * Qi_f)$</td>
<td>Species response is a function of %AES uptake in a 1km square plus a function of the quality of option implementation ($Qi_f$).</td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>$Sr_f = (\beta_1 * U_f) + (\beta_2 * Qi_f) + (\beta_3 * C_f)$</td>
<td>As above but now also testing for an independent additive function of controlling variables in the wider 1km square such as habitat composition ($C_f$).</td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>$Sr_f = (\beta_1 * U_f) + (\beta_2 * Qi_f) + (\beta_3 * C_f) + (\beta_4 * (Qi_f * C_f))$</td>
<td>As above but here the hypothesis is tested that species response is further explained by an interaction between quality of option implementation and controlling variables, for example where 1km squares of intermediate complexity with richer species spools but also potential for response to AES amplify the positive effect of high quality option implementation.</td>
<td>As above</td>
<td>As above</td>
</tr>
<tr>
<td>$Sr_f = (\beta_5 * Sr_l) + (\beta_1 * U_f) + (\beta_6 * U_l) + (\beta_2 *)$</td>
<td>As above but now including main effects of species response</td>
<td>Species abundance change, AES uptake, landscape and habitat</td>
<td>Tests the additive and interacting effects of AES</td>
</tr>
</tbody>
</table>
Table 14 GLMM model schema for detecting landscape-scale AES effects on species abundance. Note that for brevity, intercepts, random effects and residual errors have been omitted. Note that the response variable Sr_f is a measure of change in abundance from one survey to the next and so the test is that all main effects (β) are $\neq 0$.

| Qi_f) + (β 3 * C_f) + (β 7 * Qi_l) + (β 8 * C_l) + (β 4 *(Qi_f * C_f)) + (β 9 *(Qi_l * C_l)) | (Sr_l), quality of implementation (Qi_l) and controlling variables (C_l) all measured in the adjacent 8 x 1km squares so that effects of density of uptake in the wider landscape can be tested. | condition variables are introduced from wider landscape (surrounding 8 x 1km squares). The model is now able to test for landscape effects. | impacts in a focal square and in the wider landscape. |
| Sr_f = (β 5 * Sr_l) + (β 1 * U_f) + (β 2 * U_l) + (β 10 *(U_f *U_l)) + (β 11 * Qi) | Here the important extra term is a test of the interaction between uptake in the focal square and in the surrounding landscape. This tests for a non-linear density effect such that species population change is not simply the sum of a focal effect and surrounding landscape effect of AES uptake. | As above. Here the habitat condition or quality of implementation for the focal and surrounding squares are combined to reduce the number of terms in the model and giving an overall measure of implementation quality for the 9 x 1km squares. | Tests whether there is an interaction between focal and landscape AES uptake i.e. whether species abundance change is not just the sum of the effect of across focal and landscape but whether there is an additional non-linear component. For example Sr_f = 3*focal uptake + 2*landscape uptake + (1.5 * focal uptake * landscape uptake). |
5.3 Detailed review of analytical techniques for landscape-scale biodiversity data

The aim of this review is to scope analytical methods that could contribute to a new monitoring programme to assess the landscape scale impacts of the new CS AES. Each proposed analytical method is described, and any previous literature or considerations related to the target species groups is discussed.

Methods are considered in light of the distinction made by NE between agreement-level monitoring and the need to detect landscape-scale effects of AES on biodiversity:

“Landscape-scale monitoring (to be informed by this scoping study) will focus on the cumulative impacts of agri-environment agreements on biodiversity. This study will propose an analytical approach and sampling designs for monitoring the cumulative impact of AES on selected biodiversity indicators in a way that enables meaningful extrapolations from study sites that are representative of the broad landscape types in England.”

5.3.1 Generalised Linear Models (GLM) / Generalised Linear Mixed Models (GLMM)

5.3.1.1 Introduction

Almost all analytical approaches described here will be derived from a linear model framework. These models are all characterised by an assumption that a response variable (such as species abundance) can be described by a distribution with a mean and some error. The mean of the distribution is influenced by predictors (such as AES density), which are assumed to be measured without error, although in ecological data this is often not the case. Linear models have high power to detect effects of predictors on the response and are also extremely flexible. A number of adjustments can be made to account for elements of ecological data that break the assumptions of a standard linear model. One key assumption is that data points are all independent from another, which can be violated by collecting data in a time series or from spatial locations that are connected by physical or biological processes, for example when a nested sampling design is used. These violations can be accounted for by adjusting the linear model to take account of the potential dependence of observations through including a temporal covariance structure, or by using a mixed effects model to account for grouping or nesting of observations.

Linear models can also be combined to create mixture models (i.e. mixtures of multiple linear models). This is commonly required where a number of processes may be determining the response variable. Measurements of species abundance are sometimes dominated by zeroes, reflecting the large number of sites where the species was not recorded at all. This can be accounted for by considering the data as originating from two processes: one determining whether a species is present or not and the second determining, if the species is present, how many individuals there are. Two linear models, one assuming a Bernoulli or binomial distribution (for the presence component) and one assuming a Poisson distribution (for the abundance component) can be combined to estimate both components of the response (e.g. Kery 2010).
Both lack of independence of observations and zero-inflated data are likely to be properties of any data collected from a new monitoring scheme. The former will arise because sampling units for several taxa are likely to be nested within 1km squares, or groups of focal and landscape squares, and therefore there will have to be some adjustment of the model to reflect the fact that samples within squares are more likely to be similar than samples between squares. Zero inflation is likely to be a property of data collected for some taxa, especially those that are rare or rarely sampled. Therefore a combination of generalised linear, generalised linear mixed and mixture models is likely to be used to analyse any new data collected. In accordance with this, the power analysis conducted in work-package 5 (Section 4.3) uses a generalised linear model framework to estimate the number of samples required to observe effects of AES options. Most of the data likely to be collected are in the form of counts (abundance or species richness), so most models would assume data arise from a Poisson distribution, with potential to adjust for any overdispersion or to use a negative binomial distribution where the fit to a Poisson is very poor.

5.3.1.2 Examples of use in previous studies

Butterflies

The UK Butterfly Monitoring Scheme uses the TRIM method for analysis (Van Strien et al. 2004), a specific case of a generalised linear model. In this case, butterfly counts at each site in each year are assumed to follow a Poisson distribution with a log link. The model allows counts to vary with both site and year and can perform estimation in the presence of missing values, bias in sampling locations, variation in recorder effort, overdispersion of count data and temporal autocorrelation in counts. However, there are limits to how much such techniques can be relied upon to ‘correct’ for sampling design shortcomings. Logically, model sophistication is not a substitute for good data and robust sampling design since, if it were, there would be no need for data. Even with the best possible modelling approaches, biased sampling in space and time will influence the representativeness of results while low sampling effort and high variation will increase their uncertainty.

The taxon literature review identified that butterfly monitoring was most often conducted using a transect approach. Where multiple transects are surveyed per monitoring unit, it is often necessary to use a mixed model structure to account for the fact that two transects within a monitoring unit are likely to be more similar than transects in different monitoring units. For example, Rundlöf & Smith (2006) used mixed models to assess the impacts of organic farming and landscape context on butterfly richness and abundance. In this study, a paired farm sampling scheme was used with multiple transects per farm. The final model had two random effects: the farm pair and farm identity to account for the fact that farms within farm pairs were more likely to be similar than farms from different farm pairs, and transects within farms were more likely to be similar than transects between farms. By accounting correctly for these sources of variation, the effects of farming type and landscape context were identifiable.

Where the loss of spatial detail and a lower level of replication within a sampling unit are acceptable then transects may be combined. This may not be the case, for example, where different transects and transect sections are designed to represent land in option versus
counterfactual, but modelling may be more efficient if linked, nominally replicate sampling units are combined \textit{a priori}, as opposed to the links being described in the model structure.

\textbf{Pollinators}

Many studies on pollinators have used simple linear or linear mixed models; for example, Holland \textit{et al.} (2015) used a structure with transects nested on farms to assess management effects on pollinators and could account for this structure by using a mixed model.

\textbf{Birds}

For more mobile taxa, such as birds, the size of the sampling unit will be larger than that for less mobile taxa and there may be no requirement to account for nested sampling within 1 km squares, if counts are summed across multiple transects or transect segments. However, if there is a requirement to assess bird responses to within-square variation in AES implementation then a number of transects could be implemented producing a nested design. This is not the case, however, if the measure of AES uptake is taken to simply be the proportion of the square in option and the total 1km square counts as the response variable (e.g. Davey \textit{et al.} 2010a; Baker \textit{et al.} 2012). Equally, patch-level bird responses within focal squares should arguably best be investigated in a separate modelling process to that used for whole-square-level counts, because the biological processes under consideration are different (habitat selection versus population change: compare Davey \textit{et al.} 2010b or Dallimer \textit{et al.} 2010 with Baker \textit{et al.} 2012). Surveying a number of separate transects may also occur for logistical reasons if it is the best way to cover the ground during survey. Depending on whether the response variable of interest is at the whole 1km square or within 1km squares (combined across transects or considered separately), a nested analysis may or may not be required at this level. However, mixed effects models may still be required if samples are nested within NCAs. The existing GLM framework designed by Freeman & Newson (2008) has been applied to analyses of bird population data from the BBS (Section 2.5.3) in an agri-environment evaluation context (Baker \textit{et al.} 2012). This approach could be used with any future monitoring data, where population change is the response variable of interest and time series data have been collected. It uses a fixed site effect to control for local variations in baseline abundance.

\textbf{Mammals}

Depending on the method of survey, mammal samples may or may not be nested within focal squares. One sampling method that has been proposed for bats would utilise multiple samples per focal square and corresponding surrounding squares, with one sample taken from each of a location within an AES option, a counterfactual location in the focal square and a location in the surrounding squares on each occasion. Repeat sampling events would different locations of each type, so would be nested within focal squares. Squares would also be nested in NCAs and random or fixed effects controls are likely to be needed to account for square identity. Moreover each block of focal and surrounding squares would also require indexing with a random effect to account for the potential similarity in their response.
Terrestrial mammal surveys are likely to involve some sort of nested design with either multiple transects or mammal traps employed per 1 km square, or the combination of data at the square (or square and AES/non-AES) level using a binomial events/trials structure, for example.

**Amphibians and reptiles**

It is thought that amphibian or reptile responses may be highly localised, and therefore it is likely that some element of nested sampling will be included to assess AES effects at multiple spatial scales. In particular, when using artificial cover objects it is normally recommended that at least 10 are located in a focal square. Transect methods may or may not be nested, depending on the design used. Herptile data are also likely to be severely zero-inflated, due to the general rarity of many taxa and the low likelihood of observation (Durso & Seigel 2015). Generating abundance estimates for herptiles is also usually not possible (Tony Gent pers.comm.) and so data are frequently analysed as presence/absence or as a proportion e.g. of refugia occupied. It may be possible to model this with mixture models, or alternatively the probability of detection can be estimated and accounted for. If there is no *a priori* reason for modelling essentially stochastic variation between individual observations then the pragmatic approach is to combine individual data points. Because detectability of herptiles can be very low, several repeat visits are suggested for monitoring. If it can be assumed that there is no mixing of populations between visits, then observations can be summed over monitoring visits, or a maximum or average taken, with no resultant need to account for temporal autocorrelation.

**Plants**

Previous analytical approaches to the detection of AES effects on plants in the wider countryside have typically employed a linear mixed modelling approach. For example, analysis of the HLS baseline data for England contrasted species compositional indices in quadrats in HLS options with an ecologically equivalent counterfactual dataset drawn from Countryside Survey (CoS). The nesting of quadrats within CoS 1km squares, within HLS agreements and within ITE Land Class required the inclusion of random effects indexed on each of these hierarchical levels so that variance components and degrees of freedom could be correctly estimated for significance tests (Mountford *et al.* 2013). Such hierarchical arrangements are readily dealt with and are included alongside parameters for the fixed effects of interest. Fixed effects would for example include an indicator of whether a quadrat is in or out of AES option, a factor indicating option type or habitat type and other informative covariates thought likely to explain further fractions of the observed variation in the response variable.

Response variables constructed from plant species data can vary from presence/absences for individual species, species richness per quadrat or other indices such as mean Ellenberg scores. Generalised linear mixed models can readily deal with the differences in error distribution associated with these different types of response (see for example Smart *et al.* 2006). Since the focus is on rare arable plants and early successional plants in general. Thus, GLMM approaches should be able to handle the implied level of nesting of samples within arable fields and stratified by option status.
5.3.2 Structural equation models

5.3.2.1 Introduction

Structural equation models (SEM) offer a flexible approach to testing for the transmission of an effect of one variable on another via an intermediate variable. SEM can therefore be envisaged as individual regression models joined in series where a variable may be treated as a response to one set of predictors and a predictor for another set of response variables. The appeal of such models is that they reflect the way in which many ecological phenomena result from a sequence of causal relationships. Thus, SEM can be used to investigate indirect effects (Johnson et al. 1991). For example, a model predicting AES impacts on food availability could be joined to a model of food availability predicting species abundance. It would then be possible to estimate both direct and indirect (via food availability) impacts of AES on species abundance. If covariates have normal error distributions and can therefore be centred and standardized to zero mean and unit standard deviation, then estimation of indirect effects simply involves multiplying the standardized regression coefficients for each univariate model in the SEM. Non-normal data types, such as counts and categorical variables, require different but readily implemented methods (Clough 2012; Grace et al. 2012).

Defining this type of model may be beneficial when it is important to understand the mechanism of AES effects, for example, to predict impacts in unsurveyed areas or to evaluate the efficacy of measuring a proxy of species occurrence (such as food availability) instead of direct monitoring. Structural equation models could also be useful in understanding the landscape scale impacts of AES; for example, is butterfly abundance more closely related to floral resources in the focal square or in the surrounding landscape?

5.3.2.2 Examples of use in previous studies for key taxa

Butterflies

For butterflies, there is evidence that impacts of some options on adult butterfly counts are mediated by changes in the availability of floral resources (De Snoo et al. 1998; Korpela et al. 2013). Although structured equation models have not been applied in the context of agri-environment schemes, Matteson et al. (2013) used structural equation models to assess direct and indirect effects of urban development on a number of flower-visiting insects, including butterflies. They found that effects of development intensity on the richness and abundance of flower-visiting insects were mediated by changes in floral resources. There is scope to apply this method to AES monitoring to identify direct and indirect effect of options, assuming relevant mediating factors are also measured. Choice of these factors will be guided by the resources available relative to the importance of understanding and providing evidence for the mechanism whereby an option impacts a target species population. However an additional and substantial advantage of this approach arises when measurement of the rare species is costly and difficult but where the preceding variables in the causal chain are more readily measured. Thus, if an SEM model can be derived from an optimal design where uptake variables, mediating variables and target species abundances are measured then the same model can be readily used for prediction of species response in other parts of the sampling domain either in model testing.
against existing monitoring scheme data or in areas with no species monitoring, or where the species is currently absent but its potential habitat is being restored. The reliability of these predictions of AES effects across the wider countryside then becomes critically dependent upon the uncertainty around the predictions.

**Pollinators**

There is good evidence that the beneficial effects of some AES prescriptions on pollinators are mediated by increases in floral resources. For example, Pywell *et al.* (2004) showed that AES options targeted at field margins and conservation headlands increased abundance and species richness of both floral resources and bumblebees. Data were investigated in more detail to show that the abundance of adult bumblebees was linked to the abundance of floral resources. In this study, two separate models were constructed to test the effects of AES options on floral resources and the relationship between floral resources and bumblebees. Applying an SEM approach to these data would allow both models to be tested simultaneously and the identification of direct and indirect AES impacts on bumblebees. Scheper *et al.* (2015) conducted a similar study and investigated the effects of AES schemes and local and landscape scale floral resources on bees by including both variables in a multiple regression with interaction terms. Using this approach, they were able to show that local AES effects were modified by local and landscape scale floral resources. An SEM approach could extend this to quantify the extent to which AES option impacts are mediated by changes in floral resources.

**Birds**

There may be a reasonable indication that the impacts of AES on birds are related to impacts on the plant community and this could be investigated in a structural equation model framework. For example, a study of Kenyan bird diversity indicated that the diversity of woody plants mediated the impacts of environmental variables on bird richness (Kissling *et al.* 2008). Structural equation models were also used to show that variation in bird species richness along the Panama Canal was related to precipitation via plant species richness (Rompré *et al.* 2007). This approach could be applied to AES impacts by evaluating the extent to which AES option impacts are mediated by changes in e.g. seed production or invertebrate abundance, subject to vegetation or invertebrate sampling providing representative data on site-specific abundance or cover of these groups. For variables measuring plant abundance and vegetation structure, it is likely that agreement-level implementation scoring will yield useful ‘mediating’ variables. This should apply to all taxon groups where plant-related resources are an important constraint on population growth and are impacted by AES options.

**Mammals**

The mechanism of any indirect effect on mammals is likely to be via habitat provision, e.g. availability of hunting or nest sites, or via food supply. If a good mechanistic understanding of impacts on mammal species can be obtained using SEM, it may be possible to infer likely effects on mammals even where no mammal sampling has been implemented. However, this would be highly dependent on pre-existing evidence of the relationship between mammal abundance, resources and how these are expected to be impacted by relevant options.
SEM have been investigated in relation to understanding direct and indirect effects of vegetation density and prey availability on bat activity, finding that both factors only have a direct impact on bat activity when the bats are specialised to forage in open habitats (Müller et al. 2012). A similar approach could be extended to an AES context if, for example, changes in vegetation structure were related to AES option implementation.

**Amphibians and reptiles**

Mechanisms of AES impact on reptiles and amphibians are likely to be mainly via habitat creation or maintenance, including ponds, areas of bare ground and extensification. These relationships do not appear to have been investigated via SEM previously, but the method should be applicable following the caveats mentioned above. In particular, given the likely importance of habitat suitability mapping for amphibians and reptiles, it may be useful to use an approach like SEM to understand the relationships between habitat condition variables and amphibian or reptile abundance.

**Plants**

A number of studies have used SEM to investigate direct and indirect effects of global change drivers on plant species indices. For example, Sheppard (2014) estimated the direct and indirect effects of N addition on individual plant species covers via changes in soil chemistry, while Smart et al. (2014) partitioned the variation in understorey plant species richness changes in British woodlands between the direct effects of a storm event and its indirect effect via changes in canopy cover and soil pH.

Assuming a stratified random sampling design and adequate data collection for early successional plant species, it would be possible to apply SEM to test hypotheses about the role of indirect effects on either individual rare plants or aggregate indicators such as annual forb richness. Recent advances have also demonstrated how SEM can be extended to include models of variation in plant species detectability, which may be useful in facilitating a possible link with opportunistic records of rare arable plants (Maxwell et al. 2016).

### 5.3.3 Bayesian belief networks

#### 5.3.3.1 Introduction

Bayesian belief networks are a way of representing interactions between different system properties through conditional probability tables. Interactions are presented in a network structure, similar to SEM, and links between variables are assigned a probability based on empirical data or expert opinion. The strength of the approach is the ability to link variables through conditional probability, so that the impacts of changing some driving variables, such as AES implementation, can be observed in the probability of the response variables given a complex network of interactions. Because the networks are parameterised with probabilities, it is possible to derive these from either existing data or expert opinion. This means the approach is often used when empirical data are lacking. In the context of AES monitoring the approach is...
most likely to be useful in scaling up responses to unsurveyed areas, where predictions could be made of the likely impact of the AES landscape on species responses.

The drawback of this approach is that outputs can have very large uncertainty attached, especially when no empirical data are available to parameterise the probability tables, and that estimates of uncertainty for expert judgements, in particular, are highly uncertain in themselves. Novel management options and novel contexts for existing management (such as landscape-level integration) must also have uncertain effects, by definition. It will always be possible to construct a Bayesian belief network by assigning the best available estimates of probability parameters to it, but caution and ongoing validation/testing will always be required because there will always be uncertainty in some parameters and how they vary with the specific conditions in which the models are applied, which may differ from those in which the source data were collected.

5.3.3.2 Examples of use in previous studies

**Butterflies**

The key requirement of implementing Bayesian belief networks is sufficient empirical or expert knowledge on taxon ecology. This requirement is likely to be met for many butterfly species given the large existing literature. Douglas & Newton (2014) constructed Bayesian belief networks for the impacts of management on occurrence of eight taxa of conservation concern, including two butterfly species. They found there was sufficient knowledge contained in literature and expert interviews to construct highly accurate models of butterfly responses to management. Therefore it might be possible to use BBNs to predict butterfly responses in unsampled areas when empirical data are lacking by considering how known aspects of unsampled areas (such as management intensity and habitat type) moderate the expected impacts of AES options. Observations would still need to be recorded to test the claims of the BBN regarding the importance of AES option impacts and other factors on butterfly abundance.

Constructing useful BBN will be made more difficult since research on butterfly responses to AES management tends to be biased towards the adult stage, whereas larval food resources and microclimatic constraints are the limiting factors for many butterfly species. Management effects on larval abundance are well understood for a few species such as the Marsh fritillary and Brown hairstreak, but not for many others.

**Pollinators**

As far as we are aware, BBN have not yet been applied to studies of pollinator occurrence or abundance. The reasonable body of literature covering AES impacts on some pollinator taxa would suggest a BBM would be feasible, at least for bumblebees and some solitary bee species where likely foraging ranges and relationships with floral resources are known, for example. This could allow prediction of AES effects in unsurveyed areas where quantitative data are lacking. However, as for butterflies, evidence for AES management is heavily biased towards adult abundance and floral resource provision for pollinators. Hoverfly larval abundance or survival has hardly been investigated in relation to AES management, and hoverfly population growth may be limited by juvenile survival.
**Birds**

There is a wealth of knowledge about the likely impacts of management on birds which could be used to construct Bayesian belief networks. One utility of this approach is to investigate potential impacts of multiple management approaches; this was investigated in relation to bird communities in Argentina (Goijman 2014). Goijman (2014) investigated how management decisions in relation to tree management in field borders and insecticide application could influence bird occupancy while accounting for the need to maintain crop yield and farmers’ wellbeing. This approach allows not only the expected impacts of AES on biodiversity to be estimated, but also any trade-offs in relation to socioeconomic targets. A note of caution, however, is that results even for well-studied taxa are not necessarily predictable, because trial evidence for management effects may come from small-scale studies, for example, which do not scale up in practice. An example of unpredicted effects is that of wild bird seed mix in the Environmental Stewardship scheme, for which evidence-based design produced a sound prescription from field-scale studies (e.g. Henderson *et al.* 2004), but further research suggested that the timing of the management was sub-optimal (Siriwardena *et al.* 2008) and there is now evidence that effects may have become more negative over time (Siriwardena & Dadam 2015).

**Mammals**

Bayesian belief networks could be applied to assess habitat suitability for mammals in unsurveyed areas. This approach may be particularly useful for mammal species which are unlikely to be surveyed in routine monitoring. A BBN approach has been used to assess habitat suitability for a rare marsupial (Smith *et al.* 2007) to create maps of suitability given different conditions. This could be applied to AES schemes to map how the distribution of rare species might change under different options. As most UK mammals are relatively well studied there is likely to be sufficient knowledge in literature and expert opinion to parameterise networks. However, as with all applications of BBN, new observations are ultimately needed to test the predictions of each model (Marcot *et al.* 2006; Hamilton *et al.* 2015) and model predictions cannot be used as a substitute for empirical monitoring data.

**Amphibians and reptiles**

Bayesian belief networks have previously been shown to have poor predictive ability for amphibians and reptiles in other regions, due to a lack of sufficient information to parameterise networks (Tantipisanuh *et al.* 2014). However, given the small number of British reptiles there may be sufficient knowledge to parameterise networks for at least some species.

**Plants**

Given that plant species monitoring methods are relatively straightforward and have a long history of application in Britain it seems unlikely that practical difficulties will reduce the availability of data to the point where BBN is the only possible approach. However, whilst common species are readily recorded, more effort is required for rare plants and this may trade-off against the effort available for targeting other groups especially if they occur in different
habitats and geographic locations. Much is known about the ecology of rare arable plants in England and this information could be used to construct proxy indicators of the quality of habitat for supporting rare plant populations. For example Plantlife have attached scores to a range of arable plants such that the sum of these scores on a site is used to derive an index for the identification and evaluation of Important Arable Plant Areas (Byfield & Wilson 2005).

5.3.4 Informative priors

5.3.4.1 Introduction

Informative priors could be used to add information from previous studies, e.g. if an effect of AES had previously been demonstrated that information is relevant to a new analysis. They could be used in a Bayesian belief network context, or in other forms of analysis. If the prior corresponds to the posterior distribution obtained from analysing new data then high confidence in the result is obtained. However, if they do not correspond, for example if there is not enough power in the new data to inform the posterior distribution, then an intermediate distribution or one dominated by the prior may occur with high uncertainty in the outcome. Informative priors do not increase the statistical power of an analysis per se, although they are efficient in allowing use of existing knowledge, and they can be dangerous to use if statistical power is low. However, if power is high then high confidence in parameter estimates can be obtained when prior and posterior distributions overlap. Therefore caution needs to be used in the application and interpretation of any approaches using informative priors. The key requirement to use this approach is the availability of existing information about the parameter values, generally gained from the literature. As with BBN approaches, false predictive outcomes may arise if prior information is biased, unrepresentative or incorrect. Examples would include: a) drawing on experimental results where short-term, fine-scale responses do not scale up into long-term, large-scale responses; b) where a controlled experimental system does not effectively represent the conditioning effects of other factors that can result in responses outside the study system deviating from those induced experimentally; c) falsely assuming that spatial gradients in driver and response translate into equivalent temporal responses when the driver is applied.

5.3.4.2 Examples of use in previous studies

**Butterflies**

To our knowledge, no studies have so far attempted to use prior information of AES influences on butterflies, but this approach could be considered given the relatively large AES literature from which parameter estimates could be obtained, at least for local scale AES impacts.

**Pollinators**

Although it appears that informative priors have not been previously applied to studies of pollinators in relation to drivers of occurrence or abundance, there would, again, appear to be some scope to do so. As observed, there is a reasonable body of literature and, for some well-documented and common species, it may be possible to use this information to construct an
informative prior of expected AES effects at the local scale. Limited literature is available to address landscape scale impacts. However, there is a strong taxonomic bias in studies of pollinators towards common bumblebees and it is likely that prior information could be found for only a few species.

**Birds**

There do not appear to be any cases of using informative priors for bird responses to AES but the large amount of literature available could allow this if seen as applicable.

**Mammals**

It is unlikely there is sufficient information on the responses of mammals to AES to use informative priors, though this may change as ongoing research is published.

**Amphibians and reptiles**

It is unlikely there is sufficient information on the responses of herptiles to AES to use informative priors.

**Plants**

For the rare arable species considered here, there is unlikely to be sufficient information on AES responses to construct useful informative priors.

### 5.3.5 Incorporating existing datasets

#### 5.3.5.1 Introduction

In contrast to the approaches above where probabilities (Bayesian Belief Networks) or parameter estimates (Informative Priors) from existing studies or surveys were used to inform analyses of new data, this section will consider directly using the data collected in existing surveys in conjunction with new data collection. There are two main types of data collected or collated under schemes that build on the efforts of volunteer surveyors or recorders: standardised schemes, such as the Breeding Bird Survey (BBS) and Wider Countryside Butterfly Survey (WCBS), and ad-hoc, unstructured recording, such as the records collated by the Bees, Wasps and Ants Recording Society. From a funder’s perspective, the field data collection for these datasets is free, and the (significant) costs of data collation and/or survey organization are often supported by independent funding streams. However, this means that sampling for these data sets is not controlled with a view to optimizing AES monitoring. This is one of several biases likely to affect the utility of these data for AES monitoring and that, therefore, must be accounted for. These are most serious in ad hoc recording where, for example, recorders often choose where to record, leading to spatial biases, do not record with the same intensity each year, do not necessarily record all species present and may be more prone to misidentification errors. Records are also often only available at a spatial resolution (1 – 10 km²) that is too low for reliable association with AES management, and therefore may not be appropriate for capturing AES
impacts. Several new methods have been developed to take into account these effects, and are generally implemented in a Bayesian framework, often building on an occupancy model structure, hence the previous terminology “INLA and Bayesian occupancy modelling” (Isaac et al. 2014; Isaac & Pocock 2015; Woodcock et al. 2016). Thus, inference is possible although it should be emphasised that the ability of such methods to in any sense correct for the problems described above depends heavily on the nature of the biases in sampling. Whether robust imputation is possible centres on the ‘missing data mechanism’. Gelman & Hill (2007) describe these mechanisms and ways of accounting for them using Bayesian methods.

Sampling biases may also affect structured surveys, for example via low observer density in remote areas and over-sampling of land nearer to human population centres, but these biases can generally be accounted for analytically. For ongoing monitoring, a bigger concern is that, while sites can be selected for monitoring, they will not necessarily be visited by volunteers, hence long-term consistency of coverage is impossible to guarantee because of uncertain observer commitment. This means that these data sets are best suited to informing about patterns at large spatial scales, wherein site turnover is merely a source of noise. They are less useful for informing about local patterns and site level effects.

Thus, at the simplest level, existing datasets can be used to validate overall population trends, or to provide background population levels against which to assess potential impacts. However, it is also possible to investigate AES effects directly using structured data sets. The BTO/JNCC/RSPB Breeding Bird Survey and UK Butterfly Monitoring scheme have already been used to investigate agri-environment impacts (Brereton et al. 2007; Baker et al. 2012). The BBS scheme collects data at a spatial scale (1 km squares) that can be matched to option uptake, and other variables, in and around surveyed squares. It is important to note that this is a structured survey with standardized survey methods and randomized selection of survey squares, including a known, designed spatial bias in the form of a stratification for observer density that can readily be accounted for in models where appropriate. Temporal repeat surveys in the same squares allow modelling of the change in population growth rate produced by AES management (among other drivers). Similar analyses could be conducted using other schemes with similar structures, such as the WCBS. Scheme data that do not come from a randomized framework, but from surveys with a standardized protocol, such as from the UK Butterfly Monitoring Scheme or BBCT BeeWalk, can also be used in the same way, although sampling biases need to be considered, either analytically or in the interpretation of results. Thus, a similar approach has been taken with the two UK butterfly monitoring schemes (Oliver 2014). Using AES option area information it was possible to show spatial, but not temporal, differences in butterfly densities related to AES. However, the shortcomings of the different schemes at the time in conducting this type of analysis were highlighted: the WCBS provided an unbiased view of the countryside but had a very short time series whilst the UKBMS had a long time series but monitoring sites were often located in areas of initially very high quality which might not be predicted to respond to AES or have high coverage of AES management.

Unstructured data sources that rely on ad hoc, rather than standardised monitoring, (e.g. records submitted by BSBI or BWARS recorders) will be less useful because of the variable and/or unknown spatial and temporal biases involved. Records often need to be aggregated at low spatial and temporal resolution (i.e. over large units) to average over patchy recording both in
space and over time, thus limiting analytical scope. However, new analytical techniques have been developed to make better use of these data by matching them to standardised survey data that have poorer spatial coverage but higher resolution and accuracy (Pagel et al. 2014). These new techniques are based on Bayesian state-space models, an extension of the Bayesian occupancy model to include observations of two different types (from standardised surveys and ad-hoc sampling). However, to date, they have been applied only to the calculation of temporal trends in abundance, so their utility with respect to tests of management effects is untested.

Incorporating data from ongoing agreement-level monitoring can also be considered in this section, although is likely to be applicable to other methods as well (for example providing data on floral resources for structural equation models). Ideally, utilising these data would be made easier by co-location of agreement and landscape-scale monitoring schemes so that complex statistical methods will not be required to assess joint patterns. Recent examples indeed show that detection of the effects of spatially differentiated management effects can be achieved using a framework that embeds a model of the observation mechanism within a model of the fixed effects of ecological interest (Simmons 2015; Woodcock et al. 2016).

In this report, we propose sampling methods that are compatible with BBS, WCBS and that are part of the proposals for a National Pollinator Monitoring Scheme, although the methods proposed here are more intensive than those used in these schemes, in order to test effects of local and landscape AES uptake. BBS and WCBS data could thus be integrated with the proposed data that would be collected in a limited number of NCAs. This integration would most profitably involve the use of BBS and WCBS data to scale up patterns detected in these focal NCAs.

### 5.3.5.2 Examples of use in previous studies

**Butterflies**

A few studies have used information from the WCBS and the UKBMS independently to identify AES impacts, with some success (Brereton 2002; Brereton et al. 2007; Oliver 2014). As the time series from the WCBS increases, the power to detect temporal effects in the dataset will also increase.

There is also potential to combine information from multiple butterfly datasets in a quantitative manner, either combining two existing datasets or combining data from one of the national-scale monitoring datasets with field-level monitoring. Pagel et al. (2014) demonstrated this approach by combining count data from the UK Butterfly Monitoring Scheme with occurrence data from the Butterflies for the New Millennium project to predict population density of the Gatekeeper butterfly in 10 km squares with or without count data. This allowed improvement of the butterfly atlas to include estimates of population density as well as abundance. In the context of AES scheme impacts, it would be possible to extend this model with covariates, such as density of AES options, to modify the predicted population densities. The approach could also be applied to combining field-scale/agreement level data with existing datasets (Pescott et al., in prep.) with the potential to extend to include covariates considering AES impacts. However, there is a cost to this approach in terms of analytical and computational time, as the modelling approach is very
complex and computationally demanding, so it may only be suited to one-off analyses. It is also more likely to be suitable for predicting abundance of common species, where a lot of data are available on both abundance and occurrence, and for which existing monitoring under WCBS is most likely to be sufficient alone. Hence there is no guarantee that such an approach would result in robust hypothesis tests, or add significant potential inference to other data sources. In part this is because such methods have to date, emphasised the production of maps and trends as outputs, rather than a requirement to detect environmental impacts in an unbiased fashion, in the presence of other potentially confounding factors.

**Pollinators**

There is one standardised survey scheme in the UK for bumblebees (Bumblebee Conservation’s BeeWalk scheme), with a broader National Pollinator and Pollination Monitoring Framework (NPPMF) under consideration (Carvell et al., 2016). However, much current pollinator recording is *ad hoc*, and several hundred thousand verified records are collected every year. Analysis of some of these data indicates that there are sufficient data to calculate species-level trends for about 50% of pollinator taxa. As yet, there has been no attempt to use these data to identify whether trends in pollinator occurrence are related to AES impacts, although this could be made feasible by adding relevant covariates to the occupancy models already developed, even if only at coarser scales than are available for birds and butterflies (Carvell et al. 2016).

Carvell et al. (2016) found that, for pollinators other than bumblebees, no existing schemes monitor abundance, although this would probably be a key objective of any new scheme resulting from the NPPMF. As described above, it could be feasible to combine data from pollinator recording and field survey data to create landscape-scale predictions of abundance; however the approach is highly computationally demanding and requires high quality data. Again, theoretically, AES predictors could be added to this model, but this is currently untested.

**Rare section 41 invertebrates**

Many section 41 invertebrate species have restricted distributions and low abundance, making them poor candidates for a multi-taxa landscape monitoring scheme of AES interventions. As discussed at the workshop (Section 2.2) and in the pollinator and butterfly reviews above (Section 2.3), rare or restricted section 41 invertebrates may be better monitored through targeted, bespoke surveys focused at the sites on which they occur. An alternative to monitoring their abundance might be analysing changes in distribution through occupancy modelling, which uses *ad hoc* data of species presence from a range of sources (Isaac et al. 2014; Isaac & Pocock 2015). Current distribution modelling of section 41 invertebrates shows that, of the 141 section 41 invertebrate species that are targeted by the new CS AES (James Phillips pers. comm.), just 34 had distribution data that were robust enough for occupancy modelling (Charlie Outhwaite, pers. comm.), though this was not done in the context of AES interventions. Occupancy modelling of distribution data in the context of monitoring the new CS AES has drawbacks as it (i) is likely to be less sensitive than changes in abundance as a response to AES in the short to medium term, (ii) is unlikely to yield results in the timescale of a specific AES (e.g. 5 years), (iii) is untested in the context of AES uptake, (iv) is limited to species for which robust time-series of data are available and (v) may not provide information at relevant spatial scales. Nonetheless, it
is a rapidly developing field of analysis (Isaac et al. 2014; Carvell et al. 2016; Woodcock et al. 2016), has been applied to analyse dragonfly distribution changes in relation to site protected (SSSI) status (Simmons 2015), and may be an option for testing the longer term effects of AES schemes for those species that are unlikely to provide robust abundance data through field survey.

**Birds**

The UK BTO/JNCC/RSPB Breeding Bird Survey is a popular, standardised, volunteer, survey scheme, providing data on bird populations for the past 20 years. These data have already been used to identify national AES impacts on birds (Baker et al. 2012), but the scheme does not include sufficiently controlled, clustered or intensive surveys to allow monitoring of local- and landscape-scale effects within NCAs. The BBS data are at a resolution of 1 km$^2$, reflecting the fact that birds are much more mobile than other taxa. A compatible, but more intensive, method is proposed in this report to monitor the local and landscape effects of CS, so integrating the new monitoring with existing BBS data would be straightforward.

**Mammals**

Although terrestrial mammals are recorded as part of the BBS, and there are a range of variously structured or unstructured ad hoc schemes, with variable geographical coverage, there is no single nationwide mammal-specific monitoring scheme, largely because a wide range of different sampling approaches would be required, but there are individual schemes such as the National Dormouse Monitoring Programme. The National Otter Survey and a new Water vole monitoring Programme set up by PTES that can provide trend and background information. Bats are monitored by the National Bat Monitoring Programme, which involves several different types of survey focused on different bat species, while growing monitoring effort using static acoustic detectors is providing good coverage in regions where schemes have been established. The latter could inform context and scaling-up for the bat monitoring proposed in this report, because it uses the same basic method as is proposed here for intensive monitoring. Data from both terrestrial mammals and bats could be used to place AES monitoring scheme data in context and to provide guidance on survey methods. It has been possible to use the BBS mammal survey data to assess population trends for several British mammals (Wright et al. 2014), and this is currently being analysed in the context of AES scheme information by the BTO, as has been done for birds. The methodology from the National Bat Monitoring Programme was used to monitor impacts of the Tir Gofal AES on bat activity (TG report 2). Overall, however, there are fewer high quality data on mammals which will limit the ability to use data-hungry analytical approaches, such as occupancy-detection modelling or state-space modelling. Given the range of methods involved and the bespoke survey approaches proposed for new surveys, integration of mammal data may best be considered in respect of population context as opposed to directly integrated analyses, although bats may be an exception in same regions.

**Amphibians and reptiles**

Amphibians and reptiles are monitored through the National Amphibian and Reptile Recording Scheme. The scheme is new and has not yet produced estimates of population trends, nor has it
been used to look at agri-environment impacts. A power analysis conducted after the first 5 years of the scheme indicated that the power to detect temporal change was reasonable for five common taxa (over 85% power to detect a 30% population change) but insufficient for the remaining four taxa (Wilkinson & Arnell 2013). Power was tested for population change of 20% or 30% and found to be as low as 15% power to detect a 20% change in adder (*Vipera berus*) populations. Therefore it is unlikely that there is sufficient power in the scheme to detect changes due to AES for most amphibian and reptile species, although the data collected could provide valuable population context for more intensive, targeted monitoring, such as on background patterns of distribution to inform inference from zero counts, for example.

**Plants**

The highest quality existing quadrat data on common and widespread plants that can be referenced to individual habitat patches have been recorded by Countryside Survey, a professional nationwide stratified random survey at roughly decadal intervals since 1978 (Norton et al. 2012). However, the last Countryside Survey was conducted in 2007, so this information is only relevant for assessing current AES impacts where such an analysis also includes a comparison between earlier agri-environment monitoring data and previous Countryside Survey data (e.g. Mountford et al. 2013). A small-scale pilot resurvey of Countryside Survey squares is currently underway but will not yield enough data to provide a counterfactual dataset for AES impacts detection in England. A number of 1km squares across English NCAs coincide with Countryside Survey squares in which measurement of change in habitat extent and condition, focussing on soils and vegetation, has been carried out since 1990. Birds were surveyed for the 2000 CoS alone (Rhodes et al. 2015). Including these squares within a new field sampling cohort would provide a subset of squares where legacy effects of previous land-use and AES could be factored into analyses of current and medium term change. Building in scheme legacy effects would, however, depend on being able to access spatially explicit data on option uptake within each 1km square, and the local sample sizes of these squares are small, so power to detect AES effects is likely to be limited.

Plants are currently monitored by the National Plant Monitoring Scheme, a new, standardised, volunteer survey scheme. Although it is early to assess the quality and longevity of this scheme, initial uptake was high in England (http://www.npms.org.uk/sites/www.npms.org.uk/files/newsletters/NPMS__Newsletter_Spring2016.pdf) and therefore it is possible that the NPMS could provide plant records that could be combined with the proposed landscape monitoring of AES effects. Grouping NPMS quadrat data by option status would maximise the detectability of scheme effects. This would, however, require more complex sampling methods. Because of the need to ensure that survey methods are simple and attractive to volunteers, it is considered unlikely that the scheme could be altered to accommodate stratification by scheme uptake within 1km squares. This will constrain the use of NPMS data in detecting option impacts. Similar to BBS and WCBS, NPMS data could be used in the scaling up of trends detected in focal NCAs with the more intensive proposed sampling.

Ongoing national monitoring of plant populations across a wide range of management options on 180 farms under Higher Level Stewardship AES (Mountford et al., 2013 and current resurvey)
may also provide data that are relevant for scaling up from focal NCAs, although a strong focus on this data set would preclude future sampling directed at new CS agreements, specifically.

5.3.6 Habitat suitability mapping

5.3.6.1 Introduction

Habitat suitability models are used to make predictions about the occurrence of species or the habitat resources they require. In the context of understanding AES impacts, the key use of HS models will be where it is not possible to measure changes in species occurrence or abundance directly. This may happen for a number of reasons:

- species are very rare
- species were not targeted for sampling
- sampling was not conducted in the location of interest

A key requirement of this approach is that a set of models to predict occurrence or suitability for target taxa either already exists or will be built as part of the analysis. Building accurate models requires a large amount of data and this requirement may be more easily met for some species than for others. Moreover, existing models will not explicitly include scheme uptake or quality of implementation of options as explanatory variables. These effects have to be assumed to be represented by the ecological conditions represented in the model and this is inevitably a weaker approach than application of a model built from simultaneous observations of uptake, ecological change and species response. A general weakness of the approach for AES evaluation is that, by definition, it assumes that patterns identified in previous studies (which may be only local or in different regions) pertain in the region and time period of interest, with no potential to verify this assumption. Fundamentally, predictions are therefore often being made beyond the scope of the source data. Patterns identified in previous studies are also usually from analyses that link abundance of adult fauna to habitat features, which may reflect habitat selection by mobile organisms rather than evidence that particular habitat types or conditions will increase populations.

If a model is available, it will be necessary to collect data on the model predictors to be able to predict suitability. It might be possible to obtain this information from AES uptake data, agreement level monitoring or national-scale habitat datasets, but in some cases new data collection might be required in the field.

Since the focus of the project is on understanding changes in species abundance over time, models that represent spatial patterns only will be less relevant. While they can be applied, their use in this context hinges on the strong assumption that space substitutes for time (Beale & Lennon 2012). The alternative is to use dynamic models of population change but, at the species level, these may not exist, may be costly to build because of the need to measure population processes, and, even when constructed, may vary in their ability usefully to predict management impacts across heterogeneous landscapes and in the presence of multiple additional controlling variables (Crone et al. 2011). High risk therefore attaches to their construction and use, although, conceptually, they ought to provide the best simulation of future species responses to changing conditions (Evans et al. 2013).
Overall, the application of these models is best suited to making indicative predictions of the possible effects on unmonitored taxa of the changes made by AES management. The models do not represent an alternative to the collection of empirical monitoring data for the assessment of the impact of AES options.

5.3.6.2 Examples of use in previous studies

Butterflies

There is good evidence that the occurrence and abundance of adult butterflies is very responsive to habitat condition, particularly with regard to plant communities and linear features. Studies have shown that the overall abundance and species richness of butterflies can be predicted by habitat condition variables, with the advantage that these variables are faster and cheaper to record than butterfly abundances (Pywell et al. 2004). These condition variables are also those that are often modified by AES, such as nectar plant abundance and structure of linear features. However, evidence is lacking regarding the rate of population growth response to these adjustments in resources.

Habitat suitability models capturing these relationships could be applied to areas where butterfly survey data are not available. Outputs from the models could be combined with existing low resolution butterfly occurrence data (such as from national recording schemes) to refine model predictions. However, there are several potential disadvantages to this approach, for example Pywell et al. (2004) showed that models performed well for overall community measures but less well for individual butterfly species. In particular, it is difficult to apply this approach to rare species or highly mobile species, where there may be insufficient data to parameterise models. The scale of the variables used to parameterise habitat suitability models might also be important. For example, Marini et al. (2009) showed the utility of including the percentage cover of different land use types to predict butterfly species richness. They found that the local landscape scale (within 100-150m of the study site) was most important in determining species richness. Many butterfly species are highly mobile, so relationships between adult abundance and habitat condition are likely to reflect habitat selection preferences rather than demonstrate effects of particular habitats or management on population growth.

Pollinators

Given the strong observed link between presence of floral resources and adult pollinator abundance, it may be possible to model the predicted distribution of floral resources in relation to AES effects to predict cascading effects on abundance of some pollinating species, if they are limited by floral resource availability. However, many studies linking floral resources to pollinators use short-term abundance as the response variable, so do not demonstrate a sustained impact of increased floral resources on pollinator population growth. A range of plant habitat suitability models already exist as part of the MultiMOVE package (Henrys et al. 2015) and could be used to predict suitability for flowers as a result of AES options. Predicting plant habitat suitability using these models has already been trialled for the Glastir Monitoring and Evaluation Programme (Emmett et al., 2014). Occurrence of nectar providing species has been modelled for
the entirety of the UK to predict national nectar resources as part of the AgriLand project (Baude et al. 2016), using habitat associations of nectar providing species to scale up to the national scale using Land Cover Map (Morton et al. 2011). If this approach were combined with information about impacts of AES options on nectar provision and spatial information on option locations, it may be possible to adapt this methodology to predict national scale impacts of AES schemes on nectar provision for pollinators. However, prediction of nectar provision falls short of predicting impacts on population growth of individual pollinator species, because the latter may also be limited by a range of other factors.

As discussed in the species reviews, the impacts of non-floral requirements of pollinators are less frequently studied, for example provision of nest sites or larval food sources for hoverflies. Habitat mapping could be applied to predict provision of some of these resources, although there is likely to be a lack of suitable information on the spatial drivers of variation of these resources. Given the differences between pollinator taxa and species in whether adult or larval food resources limit population growth, habitat suitability modelling is not likely to be viable for a range of pollinator species.

**Birds**

A range of habitat suitability models for bird species have already been defined. Several use a GIS-based approach where known habitat requirements are identified from available GIS data and overlaid to create a map of habitat suitability. For example, the approach has been used to identify suitable breeding areas for stone curlews in Essex (Thompson et al. 2004). This approach could be used to identify parts of the landscape where agri-environment options might be predicted to have the greatest impact. However, there are drawbacks to this approach, which is only semi-objective (i.e. thresholds of suitability applied to GIS layers are often based on literature values or expert opinion).

The higher mobility of birds means that “habitats” are defined at larger scales than for most other groups, or are divided with respect to life-history function, e.g. “nesting habitat” versus “foraging habitat”. Migration provides further complications for some species, but habitat suitability models can still be constructed, including land cover variables, for example, by season or reflecting feeding and nesting habitat availability. However, key habitat variables to define “suitability” are not necessarily known or available in large-scale datasets, limiting how reliable models can be. For this reason, and the fact that birds are already the best surveyed of all taxa, there is likely to be limited scope for using habitat suitability modelling to inform about bird responses to AES. Exceptions may be where target species are too rare for abundance models to be fitted, but data on key habitat variables can be collected. Predictions of potential effects on the birds could then be made, both in areas where they are already found and in (adjacent) areas into which population spread is targeted. Note, however, that this would not necessarily reflect real population change, so the monitoring results would relate to the habitat and other predictors alone, not to the birds.

**Mammals**
Similarly to birds, some mammals are highly mobile and it may be difficult to collect sufficient data to parameterise habitat suitability models, other than via a GIS-based approach. However, it might be possible to use these relatively coarse models to assess changes in suitability due to AES-related changes in habitat extent or condition. Unlike birds, there is no national, standardised, recording for most mammals and, therefore, the ability to model habitat suitability, even in a broad-brush manner, may be useful to identify where we might expect to see changes in mammal populations. For example, a combination of national scale datasets and radio tracking was used to construct a model of grey long-eared bat occurrence at national and local scales (Razgour et al. 2011). This type of model could be applied to predict how AES-mediated changes to the habitat structure would affect this rare species, but clearly would not represent an evaluation of actual management effects. Also, habitat suitability modelling is possible for bats (Bellamy et al. 2013).

Amphibians and reptiles

For amphibians and reptiles, the existing knowledge on distributions is generally much poorer than for the other groups, therefore more insight might be gained from the use of habitat suitability models if sufficient data were available for model parameterisation and sufficiently reliable models could be constructed. Indeed, a monitoring approach building heavily on models of habitat suitability has been proposed and some work has been conducted to define a protocol for Habitat Suitability Indices for British reptiles (Brady & Phillips 2012). A metric could be based on vegetation properties such as habitat type, patch size and connectivity, which could potentially be obtained from field survey or remote sensing. In combination with an intensive field survey, the hope is that it would be possible to construct habitat suitability models which could be used in conjunction with presence/absence data obtained in a more extensive national scale field survey. One potential drawback is that it is likely that some relevant AES options will be too small to be mapped at a national scale (e.g. distribution of ponds, fine scale habitat heterogeneity) therefore it may not be possible to apply methods outside of surveyed areas. However, the method could still be used to assess predicted impacts on amphibians and reptiles within focal and landscape study squares.

While HS modelling is less developed in general for this group, recent work on modelling factors controlling Great crested newt abundance shows what can be achieved when extensive data collection and state-of-the-art modelling are combined (Bormpoudakis et al. 2016).

Plants

Habitat suitability models for British plants are available in the MultiMOVE package (Henrys et al. 2015). These models cover 881 vascular plants including 97% of CSM indicators and the dominants of most habitat types in Britain. They were built by modelling presence/absence data in terms of seven explanatory variables (mean Ellenberg fertility, mean Ellenberg pH, mean Ellenberg soil wetness, cover-weighted vegetation height, minimum January temperature, maximum July temperature and annual precipitation), using five statistical techniques to produce a small ensemble that can be used for prediction. These models are accessible in an R package available from the CEH website.
A constraint on the use of MultiMOVE is that it omits many of the rarest plants in the UK flora. A method for quantifying the association between the more common plants covered by MultiMOVE and rare plants was developed and tested by (Smart et al. 2015). Since the abiotic conditions controlling habitat suitability for common plants are estimated by MultiMOVE, this method allows rare species’ habitat suitability to be inferred from the joint presence and absence of their more common neighbours. The approach could be readily extended to rare arable plants but is currently limited by the availability of quadrat data that sample these populations and their associated species. Moreover, for rare arable plants and early successional species in general, habitat suitability alone may often not equate at all with actual presence because of the importance of specific disturbance regimes coinciding with recruitment from a patchily distributed seedbank. This means that more sophisticated models also including the latter factors in some way would be required.

5.3.7 Remote sensing

5.3.7.1 Introduction

Remote sensing could involve a number of different approaches, from radio-tracking of individual animals to use of satellite data. Here we will assume that the main interest will be in using remotely sensed imagery, radar or Lidar from satellites, planes or UAVs to assess land-use properties. Data collected in this manner could potentially inform us about:

- Location of different land use types within the landscape including properties such as area and connectivity
- Condition of habitats or other landscape features
- Impacts of AES on habitat extent or condition

Essentially, these represent approaches for identifying predictor variables for use in the habitat suitability models described above, as well as for GLMs and other approaches. Assessing change in gross habitat extent is relatively straightforward, provided that suitable earth observation (EO) data are obtained at reasonable temporal intervals, although finer habitat divisions, such as between different crops or types of grassland, are less reliable. The ability to identify change in habitat condition with EO was recently reviewed by JNCC/Defra (Gerard et al. 2015). The authors reported that while some habitat condition elements were not easily measured via EO (e.g. the presence of individual plant species), there was good potential to measure other condition attributes, such as the extent of bare ground, woody vegetation height and productivity. Whether or not these attributes could be used map habitat changes caused by AES option uptake is still to be evaluated. For example, AES extensification options could potentially lead to a decrease in NDVI-measured productivity, but this would need to be validated with field data, although hedgerow planting should be straightforward to detect. Therefore, using remotely sensed data to identify and attribute change in condition to AES may be a long term goal rather than something immediately achievable.

Newer methods for EO data collection are under rapid development but remain expensive and untested, specifically in terms of their ability to detect the level of finely resolved structural change that may result from the implementation of AES options. There is an important trade-off in the use of remotely sensed data between the resolution of information available from remote
sensing and the spatial coverage possible. Satellite information allows monitoring at a national scale but the resolution (10-20 metres for free products) is too low to pick up many AES-related changes. At the other end of the scale, UAVs can be used for very detailed habitat mapping and may be more suited to assessing changes in habitat condition, but are limited by the flight time possible, difficulties involved with data processing and the requirement for field validation; they may not offer significantly more cost-effect data collection than “live”, human observation. Currently, cost and data processing complexities mean that satellites are likely to be the major source of remote sensing data.

An essential consideration in the use of remotely sensed data is the processing required to translate the signals captured via sensors into variables used to represent landscape elements such as habitat type. Therefore, use of these data needs to rely on existing models to convert electromagnetic signals into sensible proxy variables. While there are already established algorithms to determine some landscape elements, more custom applications, such as observing changes due to AES, would require new models and ground-truthing to be conducted. Once proxy variables have been defined, data collected via remote sensing could be used in conjunction with most of the analytical approaches discussed here; for example as a predictor of habitat suitability for target taxa. The advantage to remote sensing is that data can be collected across sampled and unsampled areas, potentially allowing inferences outside the sampled NCAs. An example of the potential for deriving finely-resolved ecological maps from existing EO datasets is provided by a new hedgerow and woody linear feature map recently produced for Cornwall (Broughton 2016 unpublished report). The new map was created by subtracting a 1x1m Digital Surface Model from a Digital Terrain Model and then filtering out all tall non-hedgerow features and other features <1m (Figure 37). Such a product has obvious potential for quantifying woody linear feature length and density and hence connectivity across Cornwall and is a compelling example of how ecologically relevant information can be extracted from EO datasets. Further work would be required to produce such a map for the rest of England.
5.3.7.2 Examples of use in previous studies

**Butterflies**

Remotely sensed data have been shown to produce predictions of occurrence of some butterfly species (Luoto et al. 2002), provided that the species in question shows a reasonable degree of habitat specificity. In this example, Landsat data were used to classify habitat types and to calculate a range of habitat indices such as diversity, patch size and connectivity which were related to the distribution of a rare Finnish butterfly. A similar approach could be used in the UK utilising Land Cover Map, so long as the LCM categories are detailed enough and at the right scale accurately to reflect butterfly habitat requirements. For example, using spectral data alone, LCM2007 was able to discriminate between improved grassland and a broad category of less productive rough grassland but further discrimination to acid, calcareous and neutral grassland was estimated by reference to other information on altitude and soil type (Morton et al., 2011). Other remotely sensed data could also be useful in characterising habitat condition. Hess et al. (2013) showed that LIDAR data could be used to calculate habitat condition variables that were linked to butterfly occurrence and abundance for some species. However, many butterfly species (including several of high conservation concern) have very specific larval requirements (for particular food plant species, sward structure or microclimate), and these cannot be quantified using widely available remote sensing products.

**Pollinators**

A recent review of the utility of a variety of remote sensing approaches for monitoring pollinators and pollination concluded that there is a range of potential ways in which remote sensing could be used, ranging from monitoring habitat extent and condition to tracking individual bees (Galbraith et al. 2015). In the context of AES monitoring schemes, it might be possible to use remote sensing to identify AES-mediated changes in the landscape related to some species of pollinators in unsampled areas. This would rely on both an ability to detect changes in habitat condition or extent due to AES with remote sensing methods, and a validated link between changes in condition or extent and abundance of pollinator species from field survey data (see Habitat Suitability Modelling above). The approach may be more suited to some elements of predicting pollinator responses to AES than others. For example, evidence from the species reviews suggests that landscape complexity is important in mediating AES impacts on pollinators and it would be relatively straightforward to use land cover maps from remote sensing to explain some of the variation in detection of AES impacts on pollinator abundance. For example, the new LCM crops map (http://www.ceh.ac.uk/crops2015) could be used to control for the amount of flowering crops present in given survey locations and years. Some AES options and the resources they provide that link to pollinator responses, such as arable field margins or pollen and nectar strips, are likely to be too small or subtle to be easily identifiable by
most remote sensing methods. Field surveys would still be required, for example, to quantify floral resources.

**Birds**

Remote sensing has previously been used to measure properties of habitats related to bird species richness. For example, remote sensing can be used to assess habitat heterogeneity, such as the distribution of vegetation at different heights via LIDAR, and this has been related to bird richness (Goetz et al. 2007). Species-specific abundance and distribution data have also been modelled with respect to gross land-use by combining Land Cover Map data with national bird survey results, but always with respect to large-scale (1km² or above) patterns (Siriwardena et al. 2000; Pickett & Siriwardena 2011). While it may be possible to assess some smaller elements of bird habitats remotely (for example hedgerows used for nesting), the majority of AES options shown to be relevant for birds in the species reviews are unlikely to be observable via remote sensing (e.g. grass margins, winter food and winter stubble). Therefore any use of remote sensing would probably be limited to assessing change in habitat extent or other landscape properties such as connectivity which may be important for target bird taxa, as well as assessing habitat context, either in site selection or to provide control parameters for analyses.

**Mammals**

Examples of application of remote sensing to identify mammal habitat suitability or condition appear to be scarce (for larger or herding mammals it is possible to use remote sensing to directly observe occurrence but this is not applicable to our native mammals). It is likely that similar considerations apply to mammals as to birds; the most relevant AES options will be difficult to capture via remote sensing. Hedgerows and grass margins may be an exception; these have been shown to be important for small mammal abundance (Gelling et al. 2007; Broughton et al. 2014) and new products are currently in development to use remote sensing to quantify hedgerow density (Scholefield et al., in prep; C. Rowland pers. comm.), while grass margins are permanent features that should be readily detectable given remote sensing with resolution below 5m. In the current scoping report linear waterways were not covered, with advice from Defra at the stakeholder workshop, and so Water voles do not feature in the report.

**Amphibians and reptiles**

Remote sensing is unlikely to be particularly useful in capturing AES related changes in habitat condition for amphibians and reptiles, as many of the relevant features (for example bare patches of sand, basking areas) will be too small to be captured via remotely sensed data. Detection of pond occurrence and/or quality may be possible with high resolution data, although it may be difficult to link this to amphibian occurrence (Gómez-Rodríguez et al. 2008).

**Plants**

It is unlikely that remote sensing will offer much added value for assessing AES impacts on rare arable plants, as relevant options and changes in condition are likely to be too small to be detectable. However, annual, UK-wide changes in extent of cultivated land are now available at
2ha resolution based on a new Land Cover Map product (http://www.ceh.ac.uk/crops2015). Specific crop types are discriminated with high accuracy and so may provide useful information on the potential availability of suitable habitat across relevant parts of the English countryside.

5.3.8 Stratified random sampling

5.3.8.1 Introduction

Stratified random sampling is a method to increase representativeness of a sample to the general population by making sure that all elements (strata) have at least one sample in them. This may not happen if samples are allocated completely randomly. The number of samples per stratum is often defined as a proportion of the total contribution of that stratum to the whole population. The key requirements of this approach are that strata can be easily defined and all elements of the population (e.g. the landscape) can be assigned to a stratum. There is an assumption that variance within the stratum is smaller than variance between strata. Samples are then randomly assigned within strata, ensuring each stratum is sampled. Assuming that this approach is conducted satisfying all assumptions, it offers a powerful basis for inference to the entire population, allowing generalisation of results obtained from a small sample to the larger unsurveyed area. This approach has been used by Countryside Survey, using strata based on Land Class, to ensure that the results obtained from the sample are representative of the situation across the UK.

5.3.8.2 Examples of use in previous studies

Butterflies

The WCBS utilises stratified random sampling, based on the density of potential recorders to maximise the number of squares recorded (Brereton et al., 2011). This ensures that more squares are available where population density is higher, but that areas of low density are still sampled. The stratification can then be used to “correct” the outputs such that the more densely sampled areas do not dominate the results.

Pollinators

None of the existing (non-butterfly) pollinator national recording schemes use stratified random sampling which limits the ability of extrapolating trends to the national level. National trend estimates were, however, achieved in the paper by Baude et al. (2016), as it used floral data from Countryside Survey, which has a stratified random design.

The Agriland project, which looked at landscape scale patterns in pollinators and floral resources, outlined two different site selection methods at different spatial scales to ensure
representative cover of the UK landscapes. Although these methods do not fit into the traditional stratified random sampling framework, they may be applicable to the design of future monitoring. At the regional scale, the project identified six focal regions to maximise representation of the range of UK landscape types based on the representation of habitats and Land Classes. For every potential combination of 100 km grid squares, the distance between the cover of habitats and land classes to the overall UK cover was calculated and the combination of squares with the smallest distance was chosen to be most representative of the UK as a whole. Within focal regions, smaller sites were selected for field surveys. A new method was developed to select sites within focal areas that maximised contrasts between sites along four axes (habitat diversity, floral resource availability, insecticide loading and managed honey bee density). Using this combination of site selection techniques, it was possible both to disentangle the effects of multiple impacts on pollinators (via the survey site selection) and to ensure that results were representative of the UK as a whole (via the focal regional selection). The drawback to this approach is that it is computationally expensive to calculate the required number of combinations of sites to select the optimum.

**Birds**

The BBS uses a stratified random sampling scheme based on the number of potential observers to attempt to cover the UK in a representative manner. This is the same approach as used by the WCBS and provides a representative picture of trends in bird populations across the UK, similarly allowing correction for variations in the density of sampling sites between regions.

**Mammals**

National-scale, standardised, monitoring of mammals is conducted via the BBS, which uses stratified random sampling and therefore produces representative trends in mammal populations for the UK. Stratified surveys have also been used to estimate trends in populations of single mammal species, such as badgers (Wilson et al. 1997) and elements of the National Bat Monitoring Programme may use stratification (e.g.,NBMS -Field Survey). Other schemes such as the National Otter Survey or the National Dormouse Monitoring Programme appear not to incorporate stratification.

**Amphibians and Reptiles**

The only current nationwide, stratified, sampling scheme for amphibians or reptiles is PondNet, a new volunteer led scheme to survey ponds for taxa including (but not limited to) amphibians (Ewald et al. 2014). However, the scheme is stratified based on known species occurrence to ensure coverage of areas with known populations, not on an independent measure. Therefore results using this approach would require weighting by stratification areas to produce representative national trends.

**Plants**

The National Plant Monitoring Scheme plot selection is stratified by diversity of semi-natural habitats within 1km squares so as to ensure high coverage of semi-natural areas and good
volunteer engagement. As this stratification scheme is biased towards semi-natural areas this will have to be accounted for via weightings in the analysis (Walker et al. 2015). Countryside Survey stratification was based on assignment of 1km squares to ITE Land Classes (Bunce et al. 1996). Proportional sampling of 1km squares within each Land Class was then used to ensure a representative sample of the range of habitats across Britain.

5.3.9 Cumulative landscape effects

5.3.9.1 Introduction

Identification of cumulative landscape effects encompasses a number of different aspects. Firstly, the composition or structure of the surrounding landscape, e.g. the connectivity of habitat parcels, might modify the responses to AES options. In particular this would be expected if target taxa exploited a large range of habitats and therefore relied on other aspects of the landscape to be present or in good condition for any beneficial impact of AES to be seen. Alternatively, creation of new habitat may only be beneficial for populations if it is connected to existing habitat. Existing information in the form of land cover maps is available to characterise landscape features and these can be used as covariates in models of species response to understand dependencies on landscape composition or structure.

The presence of AES options in the surrounding landscape may affect responses to AES in focal squares. For example, if the density of AES in the surrounding landscape is relatively low, it might be predicted that AES options in the focal square may produce an attraction response, drawing individuals from the surrounding landscape into the focal square. The amount and location of AES in the landscape can be determined from existing uptake maps and included in models of species responses in the focal square as a covariate.

Additionally, the species population trends in the surrounding landscape may influence population trends, and therefore AES responses, in the focal square. For example, if the overall population trend in the landscape is negative, then it might be expected that AES responses would be identified as a less negative response, rather than an increase in population (assuming a positive impact of AES). Ideally, this effect would be picked up by choice of a suitable counterfactual. Alternatively, if population trends are increasing in the landscape, potentially due to AES implementation, it would be expected that trends would also increase in the focal square, even if no AES was in place. This might reflect either a background increasing trend or spill-over from the increasing populations in surrounding squares. To measure this effect, it would be necessary to collect information on species trends from the squares surrounding a focal square. The landscape species’ abundance could then be included as a covariate in a model of species response in the focal square. Comparisons of sites with different levels of surrounding AES management would then be required to separate the possible causes of the pattern, i.e. background changes versus spill-over.

Lastly, as a corollary to the above argument, the presence of AES in the focal square may influence population trends in the surrounding squares. Assuming AES impacts are positive, increased population size in the focal square due to high implementation of AES may lead to
increased populations in surrounding squares, even if the level of AES in surrounding squares is low, via spill-over effects. To find this effect it would also be necessary to measure species populations in surrounding squares, and in this case the landscape species abundance would be the response variable whilst AES density in the focal square is a covariate. It is perhaps worth mentioning that this mechanism is the same as this spill-over effect discussed in the paragraph above, as landscape and focal square choice is arbitrary, but the perspective has shifted towards monitoring impacts of a single area (focal square) on multiple areas (landscape squares). Assessing spill-over into multiple landscape squares also requires more care in selecting covariates, as the population trend in the wider landscape (outside of the landscape sample squares) is not known. However, this form of monitoring represents a direct test of the intended mode of action by which CS is intended to influence national populations of target taxa. 

A further implication of the arbitrary division of sampling units into focal and landscape is that their non-independence means that landscape squares will be potentially influenced by source or sink effects at work among a further outer set of surrounding squares and between the squares making up the surrounding set immediately adjacent to the focal square. So, for example, total abundance across an adjacent series of surrounding squares is more likely to be influenced by reciprocal source-sink effects than if each square were more separated in space.

Upscaling, the process of utilising observed responses in focal areas to predict responses in the wider landscape is not strictly a landscape effect but is also considered here. In this case, the model covariates are all observed in the focal or surrounding squares, but the derived model is then used in a predictive capacity in new areas. Thus, prediction and up-scaling will rely more heavily on remotely sensed data where other covariates are not available from the wider landscape. The inability of remotely sensed data to capture the fine-grained detail of option impacts will inevitably mean that up-scaled predictions based on remote-sensing and other coarsely resolved explanatory variables will be much more uncertain. Therefore a desirable feature of the analytical approach would be to quantify the loss of precision resulting from application of a model solved using progressively fewer, more coarsely resolved input variables. This, in turn, emphasises the importance at the outset of designing the most robust sampling scheme to detect landscape scale AES impacts. Such a sampling scheme should support the strongest test of the hypothesis of landscape-scale effects resulting from AES implementation, from which a predictive model can be derived. A minimum adequate model might well include agreement- and habitat-parcel-level measurements of habitat condition, as well as area of uptake and other measures of habitat area, and linear feature density and condition in the wider sampling unit. Thus, optimal prediction and upscaling would require these covariates also to be measured. However, guidance on the trade-off between model precision and not measuring the more costly covariates could be explored by inspecting model performance having set the more onerous covariates to their average values across the hypothesis-testing and model building sample. Cross-validation based on hold-out sampling would be used to achieve this sequential assessment of model simplification on prediction.
5.3.9.2 Examples of use in previous studies

**Butterflies**

Landscape scale responses to English AES have been shown for some butterfly species, both for overall AES uptake, and for combinations of specific options targeted at butterflies (Brereton 2002; Brereton et al., 2005; Oliver 2014). Butterfly species differed in whether they showed a stronger response to AES/option uptake at the 1km or 3km scale (Oliver, 2014).

Several studies mentioned above have considered other landscape scale impacts on butterflies, either through including landscape variables as model predictors (e.g. Marini et al., 2009) or by upscaling predictions to the landscape scale (e.g. Pagel et al., 2004), which could be extended to consider AES impacts. Cabeza et al. (2010) used a combination of national-scale butterfly monitoring data and regional-scale field surveys to assess conservation priorities for butterflies in Finland. Species distribution models using Generalized Additive Models were built using data from the Finnish Butterfly Inventory to identify areas of Finland with high conservation importance for butterflies and the important land-use and climatic drivers. At the regional scale, two case study areas were defined and field survey data was used to identify drivers of butterfly abundance using vegetation characteristics and land management as predictors. This combined approach efficiently makes use of existing data to look at drivers at multiple spatial scales and could be extended to consider AES options as drivers. However, in this study there was no attempt to assess landscape scale effects quantitatively, i.e. the national and regional scales were considered separately.

**Pollinators**

There is good evidence that the composition of the surrounding landscape influences the effect of local AES schemes on pollinators. This has been demonstrated in studies such as Scheper et al. (2015) who investigated how both local- and landscape-scale effects of floral resources modified AES impacts on bees. There is also some evidence that habitat complexity influences the benefit obtained from AES, with low-intermediate complexity habitats benefiting the most (Scheper et al., 2013). Habitat complexity could be determined by remote sensing approaches and mapped to identify areas where AES options may have most impact, although, as noted above, many key landscape features for pollinators (e.g. uncropped field margins) are likely to be too small to be easily identifiable by most remote sensing methods.

**Birds**

Landscape scale responses to AES have been shown for birds in one specific context. Baker et al. (2012) demonstrated that, for some options that act on birds in winter, larger responses were seen in terms of bird population growth rate for some species when management in the surrounding 25 km² was considered compared to the only considering management in the 1km survey square. This shows the variation in patterns of inter-seasonal movement between species and how this affects the scale at which habitat must be considered to represent the resource base used by birds in a given focal area. There is little evidence that surrounding landscape complexity impacts bird abundance or species richness, potentially because high mobility means that birds can choose where to forage within the landscape (Batary et al., 2011), although this may be an artefact of the
scale at which “local” and “landscape” habitats have been considered in multi-taxa studies. This high mobility may also limit the detection of spill-over effects, but the influence of a given management patch is still likely to be larger closer to that patch than further away, so the key to the detection of spill-over will be the sampling of unmanaged areas within the potential sphere of influence of focal management, given that the biological mechanism will involve either the settlement to breed of “extra” adult birds whose survival has been made possible by winter options, or the dispersal of “extra” juveniles raised in managed areas to breed elsewhere. Specific spill-over effects from local AES management or particular agreements have yet to be demonstrated.

**Mammals**

One of the most obvious potential landscape effects on mammals is that of habitat connectivity. Connectivity is known to be important for a range of mammals using farmland (Gelling *et al.*, 2007) and is likely to modify responses to AES. In addition, small-scale spill-over effects have been suggested for small mammals, with field margin treatments apparently having a positive effect on mammal abundance in nearby non-AES margins in one study (Broughton *et al.*, 2014). Again, the size of the relevant landscape is likely to be variable between mammal taxa.

**Amphibians and reptiles**

The structure and size of the relevant AES landscape is likely to differ between amphibians and reptiles. Amphibians are more likely to be influenced by the density of surrounding ponds, whereas reptiles may be influenced by terrestrial habitat heterogeneity and connectivity. Compared to mammals and birds, the size of the landscape is likely to be small, because most amphibians and reptiles will have dispersal distances of less than 1 km and therefore the scale at which it is meaningful to infer landscape effects for these taxa will be smaller.

**Plants**

Plants are sessile organisms and therefore rely on dispersal via seeds, spores and vegetative fragments. Once dispersed into a patch, environmental filters operate to increase or to decrease the likelihood of establishment and persistence. Being mostly annual species, rare arable plants, and early successional species in general, are inherently capable of rapid population dynamics in response to environmental changes, so have a high capacity to respond to AES intervention. However, their dependence on ephemeral gaps or appropriate cropping regimes and fertility result in a strategy that sees many species persisting in patchily distributed seedbanks whilst remaining very rare and localised above ground, especially so in currently unsuitable habitats. These species typically trade-off high biomass accumulation with rapid growth and investment into seed. Hence, even if present as adult plants, populations are often transient and small. All these factors conspire against easy detection using standard vegetation monitoring methods and, if they are to be recorded during survey, require precise targeting on known extant populations and their surrounding areas of habitat. The requirement for specialised combinations of
disturbance and edaphic conditions also suggest that landscape-scale spill-over effects are only likely to occur if dispersal in space is accompanied by an increase in availability of appropriate gaps via management.

5.4 General issues identified by analytical scoping

5.4.1 Taxonomic level of interest

For some analyses it may only be possible to look at combinations of multiple species as a response variable due to the rarity of individual taxa, so analyses may be restricted to overall community indices plus individual analysis of the most commonly recorded species. This raises the question of where we can assume that species respond similarly and where can we not make this assumption? If taxa are grouped that respond very differently to AES, then we may obscure true effects. A general recommendation would be to analyse at the species level where possible, then use scientific knowledge to group species into potential “functional response groups” of taxa that are predicted to respond to AES in similar ways. If successful, this approach would allow identification of effects not obvious when all taxa are combined, but the potential biases in grouping taxa must also be understood.

5.4.2 Rarity

Most species are rare to some degree and we are often most interested in the rarest of these. For example there is, as would be expected, an interest in how AES can contribute towards conservation of our rarest species, those designated in Section 41. However, the likelihood of actually sampling a Section 41 species in a field survey targeted at a more generic species group is very low, meaning it is unlikely that sufficient information would be gathered on the ground to assess AES impacts on rare Section 41 taxa. It may be tempting therefore to use approaches such as habitat suitability modelling and Bayesian Belief Networks, which rely on previously published literature or expert knowledge. However, there is a danger to this in that both approaches rely on high quality existing information which, if lacking, could bias our interpretation of likely AES impacts on very rare taxa. Therefore, to understand responses of the rarest species correctly it is likely that some targeted surveying would be required which would be separate to the main body of monitoring that will focus on widespread species. It is highly unlikely that a single sampling framework could adequately cover multiple rare species as well as the less restricted taxa and habitats that inform about the condition of the wider countryside. It may also be more appropriate to conduct surveys for rare species at the agreement scale, as very rare species may be unlikely to show observable landscape-scale responses to AES implementation.

5.4.3 Legacy effects

For the majority of new CS agreements, it will not be the first time an AES has been implemented in that location. There are likely to be continuing effects from previous schemes
such as old CS, ELS and HLS, even when no scheme has been applied in recent years. For example, creation of a wildlife-rich habitat as part of a previous HLS agreement will mean that a new agreement could start out with a higher ecological condition than surrounding non-agreement areas, separate from the impact of new options. It will be important to account for these effects by including them as covariates in an analysis to assess the impact of the new CS scheme properly, requiring high quality data on the location and duration of management under previous agreements. If spatially referenced data are available on the entry date of each parcel into AES management, this will facilitate the inclusion of legacy effects.

5.4.4 Spatial vs. temporal analysis

One of the ultimate aims of any AES monitoring scheme is to assess the long-term effects of AES on population trajectories of individual species. However, this relies both on reliable long term population data for the species of interest and on long-term implementation and monitoring of AES. Often, these data are not available, either because the population trends are not available for taxa, or because the AES scheme in question has not been running for long enough to find a relationship. Year-to-year fluctuations in other variables such as precipitation can also obscure identification of AES impacts. On the other hand, spatial analyses can be conducted almost immediately after scheme implementation, although caution is required in interpreting results, particularly where legacy effects may be important and with mobile taxa for which short term abundance may reflect habitat preference. The only requirement is to sample areas along a gradient of AES and this does not necessarily have to be repeated. However, it is not possible to directly infer AES impacts on population growth from a purely spatial analysis, therefore some degree of temporal replication is required if this is the aim of the monitoring scheme. It is also impossible to separate effects of location bias from those of scheme management (e.g. the possibility that management is located on previously above-average sites): with an appropriate sampling structure, legacy effects of different types and durations could indicate the extent of AES impacts.

5.4.5 Extension to non-target taxa

Some elements of the proposed analytical workflow for butterflies will also apply to moths (non-target taxa). However, generally the availability of existing datasets is lower for moths and they tend to be recorded with different methodology, so they are likely to require a customised analysis. AES impacts have been shown for moths, even when they were not the target of the options (Merckx et al. 2010b).

In general, the survey requirements of many different taxa are so variable that it is unlikely that surveys designed for one group will provide much field data for non-target taxa to use in analysis. Therefore any extension to non-target taxa would probably be via add-on modules or by using existing datasets.

5.4.6 New types of data

153
One new type of monitoring that has been developed is to use genetic methods of identification, e.g. environmental DNA and meta-barcoding approaches. We have not covered this type of data as part of this review as the analytical strategies to analyse this highly multidimensional data will be very different and few of the methods discussed here are likely to apply. This approach currently also only allows consideration of presence/absence (and, perhaps, biomass) so it is not likely to provide sensitive evaluation of effects on species’ abundances.

5.5 Consideration of analytical power

Simple generalised linear models will have the greatest power to detect effects due to new CS. The representativeness of any results to the UK as a whole would be increased by a stratified random sampling scheme. Other methods will not directly improve statistical power to detect effects, but may allow extension to prediction outside of the survey square (e.g. habitat suitability mapping) or increase understanding through the combination of multiple datasets. Further consideration of power is made in the power analyses in work package 5 (Section 4).

It might be possible to increase power to detect marginal effects by using Generalised Estimating Equations. This analytical method can have high power to detect effects across a whole population (e.g. to estimate the mean impact of AES on bird abundance across the sampled NCAs) but cannot give any information on conditional responses (such as how the AES response differs between NCAs). The method is therefore only suited to looking at effects across the whole population of samples and may be unsuited to a survey design where samples are nested and there is some interest in how the responses differ between nests (i.e. NCAs).

5.6 Analytical approaches in relation to different survey designs

The majority of analytical approaches outlined will be applicable to a number of survey designs, e.g. different methods of species monitoring and different locations of samples. Most likely types of collection (counts of observed species along transects, trapping, observation of mammal signs per sample unit) will produce count data which can be analysed with the exponential family of distributions accommodating zero-inflation and overdispersion where necessary. From an analytical point of view, the monitoring methodology may have little impact on the analytical protocol. However, the choice will affect interpretation of results in a landscape context, for example pan traps may attract pollinators from a wider area than is sampled by transects, and would provide data to species level resolution.

Using a stratified random sampling scheme to select sample locations will allow conclusions to be made regarding the whole of England, but only quantitative variables can be used to stratify in this way. This is a key consideration when sampling within NCAs: if NCAs (a categorical variable) were used to stratify the samples it would be necessary to sample all NCAs, which is not feasible. An alternative would be to stratify by characteristics such as land class which are common between several NCAs. Alternatively, the method used to select regions for analysis in
the Agriland project, where the set of regions covering the greatest breadth of environmental conditions was chosen, could also be used to provide representative results.

Co-location of measurements will not increase power to detect AES effects on single taxa but will allow effects to be assessed across taxa, as well as interactions between taxa. For example, in a structural equation model approach it may be useful to co-locate agreement level monitoring of plants with field surveys of pollinators, or to consider the attraction and possible influence of predators in models considering prey taxa. It may be that the level of co-location is at the 1 km square, as anything below that is difficult logistically, but even at this scale useful questions could be asked by combining two co-located datasets.

5.7 Detailed review of analytical techniques: conclusion and comments

Any analytical protocol for analysing data from a new monitoring scheme is likely to be built on a generalised linear model framework. It is almost certain that some level of nesting will apply, so mixed effects models will be used, although it may prove more efficient to pool data and to use a simpler model, and in some cases it will be important to account for other elements of the data such as zero-inflation. Co-locating measurements across taxa will allow more complex and ecologically informative analyses to be conducted.

At a national level, existing remote sensing data is likely to be used to some degree, as a minimum to quantify habitat types in focal and landscape squares, and across NCAs. Use of remote sensing data in conjunction with habitat suitability models is also a promising approach to predict AES responses in unsurveyed areas, i.e. enhancing the information extractable from vegetation and habitat monitoring, but requires further testing and field validation. Although these models may be imprecise, they may be useful to target AES monitoring efforts or to predict national scale scheme impacts. These models could also be informed by structural equation models to investigate mechanisms of AES impact and by existing datasets from volunteer schemes.

Methods that utilise expert or prior knowledge (Bayesian belief networks and prior information models) are the most risky because they do not involve analysis of new observations of ecological change, but instead use expert knowledge or quantitative information from previous analyses. Therefore, they are only recommended as an aid to estimating impacts when information is absent or sparse; they are not a valid substitute for the collection of empirical monitoring data. Without new observations made to test the predictive claims of such models they cannot be considered to yield robust evidence for AES impacts on species’ abundance.


158


Hails, R.S., Freeman, S., Henrys, P. & Siriwardena, G. (2013). Determining and increasing the sensitivity of existing environmental surveillance monitoring networks to detect unanticipated effects that may occur in the environment in response to the cultivation of genetically modified crops. Defra research report number CB0304


161


Macleod, A. (1999). Attraction and retention of *Episyrphus balteatus* DeGeer (Diptera: Syrphidae) at an arable field margin with rich and poor floral resources. *Agriculture, Ecosystems & Environment*, 73, 237-244.


