JNCC Report 756

Review of monitoring biodiversity effectively at differing scales
(Guidance report)

Harris, M. and Hoskins, H.

January 2024

ISSN 0963 8091

JNCC's report series serves as a record of the work undertaken or commissioned by JNCC. The series also helps us to share, and promote the use of, our work and to develop future collaborations.

For further information on JNCC's report series please contact:
Joint Nature Conservation Committee, Quay House, 2 East Station Road, Fletton Quays, Peterborough PE2 8YY
https://jncc.gov.uk/
Communications@jncc.gov.uk

This report was produced by JNCC for Defra through the Natural Capital and Ecosystem Assessment (NCEA) programme.

This report should be cited as:
Harris, M. & Hoskins, H. 2024. Review of monitoring biodiversity effectively at differing scales. *JNCC Report 756 (Guidance Report)*, JNCC, Peterborough, ISSN 0963-8091. https://hub.jncc.gov.uk/assets/145a3536-acae-4fad-8242-85b52cfe8c05.

Acknowledgments:
This project was funded by the Department for Environment, Food and Rural Affairs (Defra) as part of the Natural Capital and Ecosystem Assessment (NCEA) programme. The NCEA programme is leading the way in supporting Government ambition to integrate natural capital approaches into decision making. Find out more at https://www.gov.uk/government/publications/natural-capital-and-ecosystem-assessment-programme.

We thank Nick Marriner from the Chilterns Conservation Board for allowing us to interview him in detail to produce the case study Appendix to this report about the 'Tracking the Impact' project and for his review of the report. We also thank UKCEH, BTO and Butterfly Conservation for their contributions to reviewing the report.

Evidence Quality Assurance:
This report is compliant with JNCC's Evidence Quality Assurance Policy https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

Summary

This report aims to provide guidance for developing coordinated, multi-scale biodiversity monitoring (summarised in Figure 1), based on a short literature review assessing approaches that have been undertaken previously.

Multi-scale biodiversity monitoring is important because:

- Ecological processes and pressures on biodiversity take place across different scales.
- Decision making takes place across different scales.
- A framework that allowed for integration of data collected across different scales could mean an increase in the sample size that can be used, without increasing the time or resource required for monitoring to be carried out.

A multi-scale understanding would therefore lead to increased ability to make robust recommendations for policy and conservation action.

The key challenges that multi-scale monitoring faces include:

- Sampling design is typically targeted at one scale, with more intensive sampling for smaller scale monitoring that would be resource intensive to replicate at large scales.
- There are often inconsistencies in protocols used across different scales, meaning that results produced are not comparable or interoperable.
- Monitoring is typically designed to answer a specific question. Questions typically relate to a single scale.
- Local priorities and contexts often determine what data are collected at a given site. This may differ from larger scale priorities and contexts.
- Whilst volunteers collect data for a variety of reasons, these are often linked to their sense of place and their drive to contribute to local conservation priorities rather than motivations that span scales.

Key activities that could be undertaken to overcome such challenges (Table 1) include:

- Ensuring interoperability between monitoring across initiatives and scales.
- Use of analytical techniques to improve interoperability even when different protocols are used.
- Ensuring complementarity, collaboration and coordination across initiatives and scales.
- Providing feedback to volunteers about how the data is used across scales to increase motivation and understanding of the importance.
- Understanding data requirements and targeting monitoring in order to combine requirements and priorities from across different scales to optimise data collection that can be of most use overall, instead of for answering one particular question.

The report’s key conclusion is that whilst there are substantial challenges to multi-scale monitoring and to combining data from multiple recording initiatives to achieve common aims, it is also an area that shows significant potential for improving environmental recording, and for increasing the applicability and efficiency of data collected. Following the guidance provided in this report will help to break down current barriers to multi-scale monitoring and to inform conservation across multiple scales in the future.

Details of an example project ('Tracking the Impact', in the Chilterns Area of Outstanding Natural Beauty) that has made use of multi-scale monitoring is provided in Appendix 1.

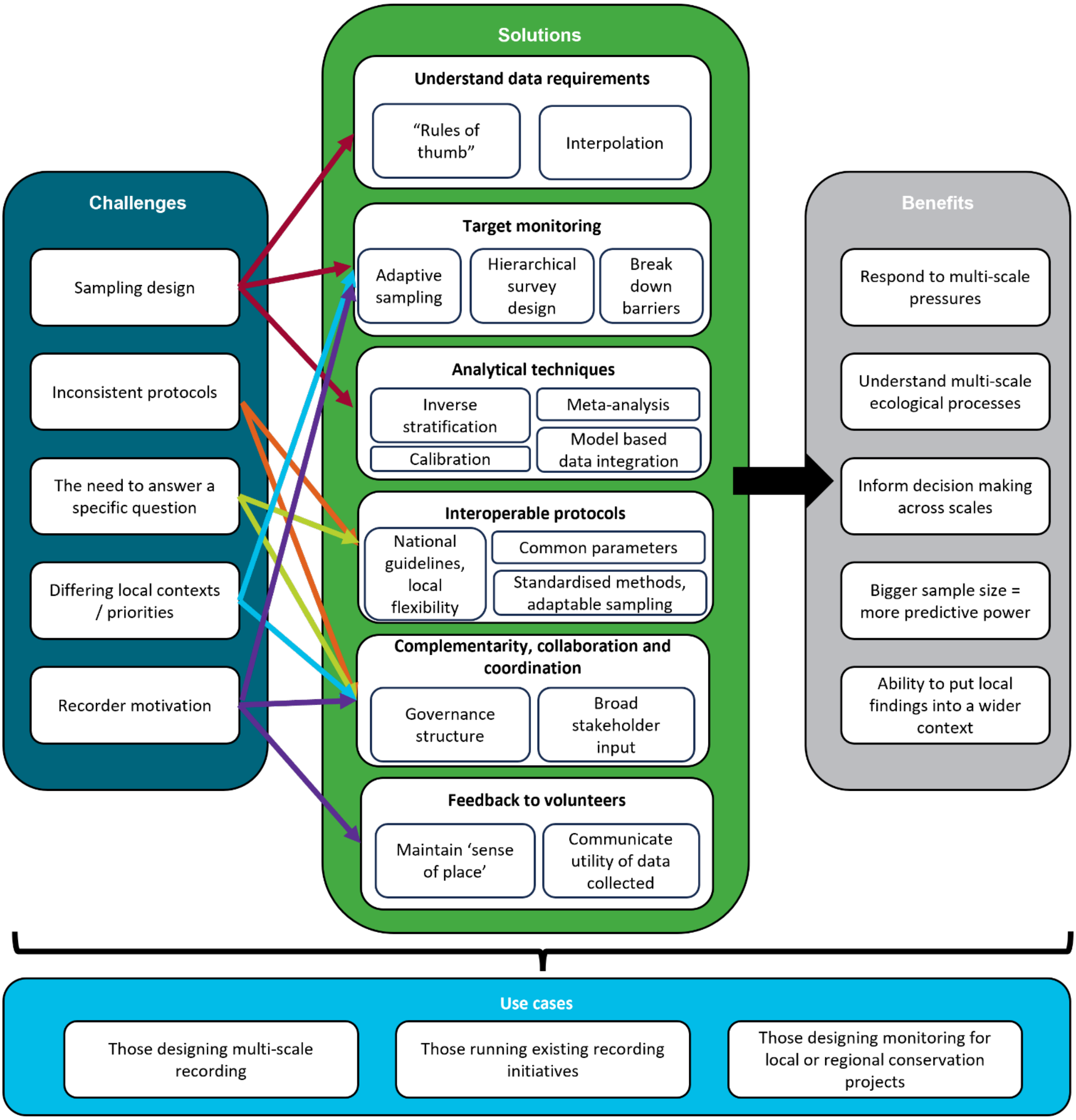

Figure 1: A summary of the challenges, solutions, benefits and use cases relating to multi-scale monitoring.

Table 1: A summary of the resources identified as part of this review that may assist with the implementation of each solution suggested.

Theme	Option	Resource and link
Understanding data requirements (Section 3.1)	"Rules of thumb"	• Border *et al.* 2019 • Pocock *et al.* 2019
	Interpolation	• Henry *et al.* 2008 • Comber & Zeng 2019 • Young *et al.* 2012
Targeted monitoring (Section 3.2)	Adaptive sampling	• The DECIDE project • Pocock *et al.* 2023
	Hierarchical survey design	• Reichart *et al.* 2021
	Break down barriers to monitoring	• Border *et al.* 2019 • Reichert *et al.* 2021 • Thomsen & Willerslev 2015 • Sparrow *et al.* 2020 • Kelly 2008 • Sugai *et al.* 2019
Analytical techniques (Section 3.3)	Inverse stratification	• Henry *et al.* 2008
	Calibration	• Henry *et al.* 2008
	Meta-analysis	• Henry *et al.* 2008
	Model based data integration	• Mancini *et al.* 2022 • Adams & Muths 2019 • Kühl *et al.* 2020 • Border *et al.* 2019 • Jarvis *et al.* 2021 • Pescott *et al.* 2015 • Henry *et al.* 2008
Interoperable protocols (Section 3.4)	National guidelines, local flexibility	• Adams & Muths 2019 • Danielsen *et al.* 2005 • Kühl *et al.* 2020 • Thomaes *et al.* 2021 • Reichert *et al.* 2021
	Common parameters	• Adams & Muths 2019 • Danielsen *et al.* 2005

Theme	Option	Resource and link
Interoperable protocols (Section 3.4) (continued)	Standardised methods, adaptable sampling	• Appendix 1 • Stauffer *et al.* 2022
Complementarity, collaboration, and coordination (Section 3.5)	Governance structure	• Kühl *et al.* 2020 • Pavlacky Jr *et al.* 2017 • Reichert *et al.* 2021 • Tulloch *et al.* 2013 • Henry *et al.* 2008 • Thomaes *et al.* 2021
	Broad stakeholder input	• Sparrow *et al.* 2020 • Kühl *et al.* 2020 • Reichert *et al.* 2021 • Pavlacky Jr *et al.* 2017
Feedback to volunteers (Section 3.6)	Maintenance of sense of place	• Appendix 1
	Communicate utility of data collected	• Tulloch *et al.* 2013 • Reichert *et al.* 2021 • Thomaes *et al.* 2021

Contents

1. Introduction

1.1 Rationale

Biodiversity is declining rapidly (Cowie *et al.* 2022; IPBES 2019; Ripple *et al.* 2017; WWF 2022). Biodiversity monitoring is essential to addressing this problem, from understanding large-scale trends, to informing and evaluating conservation actions on-the-ground (Kühl *et al.* 2020; Navarro *et al.* 2017; Niemelä 2000). Monitoring can also provide a platform for the public to engage with and learn about nature through citizen science programmes (Turrini *et al.* 2018).

Currently, most biodiversity monitoring takes place at a single spatial scale, with little integration across scales (Henry *et al.* 2008; Reichert *et al.* 2021; Stauffer *et al.* 2022). For example, in the UK there are a number of national monitoring schemes focusing on specific taxa, such as the Wider Countryside Butterfly Survey (WCBS), the Breeding Bird Survey (BBS) and the National Plant Monitoring Scheme (NPMS). These are largely based on dispersed stratified random sampling across the whole country. There is also widespread local monitoring, such as local natural history groups, assessments of SSSIs (Sites of Special Scientific Interest) and evaluations of specific conservation action related to individual projects. However, data flow between the two scales is minimal, with the data from local monitoring being used primarily for local applications and the data from national schemes being used primarily at a national scale (noting that there are exceptions).

Given that pressures on biodiversity take place across scales, multi-scale monitoring is necessary to target effective interventions and to avoid relying on flawed assumptions that ecological processes take place consistently across scales (Ascensão *et al.* 2018; Uchida *et al.* 2021). This lack of data flow is therefore problematic. Multi-scale monitoring could help to identify links between local conservation efforts and large scale population trends (Pavlacky Jr *et al.* 2017). Decision making also takes place across different scales, so having data available that allows for consistent and comparable analysis in the context of each scale would be useful from a policy and land manager's perspective as well (Paloniemi *et al.* 2012). A framework that allowed for integration of data collected across different scales could also mean an increase in the sample size that can be used, without increasing the time or resource required for monitoring to be carried out (Reichert *et al.* 2021). This would lead to increased predictive power and inference, and so a better understanding of complex environmental problems, and an ability to make more robust recommendations for policy and conservation action (Danielsen *et al.* 2005; Henry *et al.* 2008).

Even when answering local questions, an ability to put findings into a wider context is useful. For example, if local monitoring is showing increased populations in an area where conservation actions are taking place, but wider landscape-scale monitoring shows that populations are decreasing in nearby areas, it could be unclear whether the intervention is attracting mobile species from the surrounding region, rather than leading to increased populations *per se*, or whether higher local productivity in this area would result in spill over to the surrounding habitat (Staley *et al.* 2021). If a species is locally increasing but nationally declining, it may be of high interest and priority to understand the trends and their reasons in more detail to inform conservation action.

Developing multi-scale monitoring is challenging (see Section 2). There have been calls for greater efforts for collaboration and communication between organisations collecting data, and research into novel methods that would facilitate the process (Danielsen *et al.* 2005; Reichert *et al.* 2021). This report explores potential solutions to help break down the barriers associated with multi-scale monitoring and meet this need.

1.2 Policy context

Gaining a more robust understanding of biodiversity and environmental change is crucial to several key international and UK policy areas. Multi-scale monitoring could increase this understanding and make an important contribution to such areas. For example, the Convention on Biological Diversity's Kunming-Montreal Global biodiversity framework includes a target to "Ensure that the best available data, information and knowledge, are accessible to decision makers." The UK Government's 25 Year Environment Plan aims to "Improve monitoring and evaluation of policies so that both costs and benefits can be more accurately estimated in future analysis," and its recent update (the Environmental Improvement Plan) continues a commitment to "monitoring of progress toward the ambitions and goals" that it sets. To support this, the UK Government's Natural Capital and Ecosystem Assessment (NCEA) programme aims "to collect data on the extent, condition and change over time of England's ecosystems and natural capital". The terrestrial NCEA programme will be running several local pilots to understand potential mechanisms and needs for coordinating multi-scale field data collection. This will be based around four Nature Recovery Projects which form part of the Natural England Nature Recovery Network. The NCEA programme is also running several national monitoring programmes such as the England Ecosystem Survey, so ensuring that both local and national cases can most effectively complement each other will be of relevance here.

1.3 Aims, scope and audience

This report aims to provide guidance for developing coordinated, multi-scale biodiversity monitoring, based on approaches that have been undertaken previously. It begins by highlighting the challenges that such a programme would face (Section 2), then explores guidance that could be followed to overcome such challenges (Section 3), and finally provides advice for several cases in which the guidance may be relevant (Section 4). It is based on a time-restricted review of the scientific literature, and so does not intend to be comprehensive or systematic. The report is considering biodiversity monitoring, but concepts are likely applicable to wider environmental (or other) monitoring contexts.

It is hoped that the report will be useful for those interested in biodiversity monitoring at multiple scales, including those designing multi-scale recording from scratch, those involved in existing biodiversity monitoring who wish to increase the applicability of the data they collect to gain a greater understanding across scales, or those designing and monitoring local or regional projects who wish to contextualise their findings with a broader spatial understanding. This could include those running biodiversity monitoring initiatives of any type, national and local government bodies, NGOs, land managers, local interest groups and citizen scientists.

It should be noted that the challenges, and therefore guidance, associated with monitoring across scales will vary depending on the species studied. For example, monitoring of a species with a large range may provide different challenges to those restricted to a smaller spatial scale in the first place. This report does not focus on any one species but aims to draw general conclusions that are likely to be widely applicable across any taxa.

The report builds on and takes into account previous work undertaken through the Terrestrial Surveillance Development and Analysis (TSDA) Partnership comprised of JNCC, UKCEH and BTO, working with partners of the Terrestrial Evidence Partnership of Partnerships (TEPoP).

2. Challenges to monitoring across multiple scales

2.1 Sampling design

One key challenge to monitoring biodiversity across multiple scales is that it requires a sampling design that would allow for effective analysis at each of these scales (Peters *et al.* 2014). Typically, however, initiatives that are aiming to collect data at a national scale do so too sparsely for the information to be of use to answering local questions. Collecting national scale data at a high enough resolution to also be useful locally would be resource intensive, requiring a high density of surveys to be undertaken across the country, and therefore large amounts of volunteer time, which is unlikely to be possible. Meanwhile, data collected at a local scale can bias results if fed into structured national scale monitoring, as it means there is a higher density of data collected from certain locations. These areas may be of specific conservation interest and therefore not representative of the wider landscape.

Sampling designs also differ across initiatives, even at the same spatial scale. Some recording asks for records from anywhere (unstructured, or *ad hoc* recording), whereas other monitoring sends recorders to specific locations to get an unbiased spatial assessment (structured recording). Unstructured recording often generates larger volumes of data than structured recording, while structured recording is a more statistically rigorous approach. Combining data from both approaches could go some way to contributing to the increased survey density required for national scale data to be useful locally, at least in locations where unstructured recording is popular. However, this presents further challenges as it is difficult to do so without undermining the statistical power of the structured surveys (Mancini *et al.* 2022).

Multi-scale monitoring would therefore require a good understanding of the data requirements at each scale (see Section 3.1), targeted monitoring (see Section 3.2), and/or novel analytical techniques (see Section 3.3).

2.2 Inconsistent protocols

Another key challenge is the difficulty in combining data from multiple monitoring programmes due to inconsistency in the protocols used (Peters *et al.* 2014; Shepherd *et al.* 2015). Current monitoring consists of a wide variety of different initiatives, each of which requires recorders to measure different factors in different ways. Initiatives vary in the taxa they focus on, the geographic scope they cover and the definitions that they use for key variables recorded. For some, a record may simply be whether the species or habitat is recorded as present at a particular location, while others may also require demographic (e.g. counts, males/females), meteorological (e.g. weather conditions) and other types of information (e.g. estimated distance from a transect line to correct for observer bias). Some initiatives have a detailed protocol to follow, and associated guidance and training, whereas others leave much of the ‘how’ up to the recorder and simply ask for submissions of records of any species from anywhere. Differences in how data are recorded and stored can also lead to difficulties combining data from different initiatives, even if similar field protocols have been followed. From a statistical perspective, the ideal multi-scale monitoring programme would follow an identical protocol at each site.

Monitoring across multiple scales would therefore require interoperable protocols (see Section 3.4) and complementarity between local and national recording schemes (see Section 3.5).

2.3 The need for monitoring to answer a specific question

Another barrier is that monitoring is most effective when it aims to answer a specific question (Pescott *et al.* 2015; Stauffer *et al.* 2022; Tulloch *et al.* 2013). For example, a national project may be interested in understanding population trends of a particular species. The information collected would need to be representative of the country. This could help understand the broad picture across the whole country to help prioritise whether this species needs to be a conservation priority based on overall trends. However, it would provide limited information about the potential drivers of this change at a local scale. Jarvis *et al*. (2021) found that national scale citizen science monitoring could not be used to detect changes in populations in areas of different agri-environment scheme practices. Monitoring aimed at a national scale is therefore not likely to meet local scale monitoring needs.

Meanwhile, a local monitoring project may investigate how a particular conservation action at a local site has influenced the populations of this species. This could help inform them on whether to continue this conservation action and replicate it in other areas. The information collected would therefore be designed around this specific question. No population data would be recorded without also recording data on the management in the area, perhaps making use of a control site as part of the experimental design. Monitoring aimed at a local scale is therefore also unlikely to meet national scale monitoring needs.

Whilst the overall results of the first example can help inform the priority of undertaking the second example, the raw data collected from each would not be relevant to the other even though data are being collected on the same species. The need to answer a specific question means that the most useful data to collect for one of the questions (representative samples across the whole country) is not the most useful data to collect for the other question (which specifically requires samples of atypical land that is undergoing targeted or no management).

However, if multi-scale monitoring were able to encourage collection of data that were most able to answer a wider variety of questions, for example through encouraging interoperable protocols (see Section 3.4) and ensuring complementarity between local and national schemes (see Section 3.5), this problem could be reduced.

2.4 Differing local priorities and contexts

Local priorities and contexts often determine what data are collected at a given site. For example, different sites are likely to have different rare species present, which may be seen as priorities for monitoring (see Appendix 1). This presents a challenge when attempting to scale up and use the data to inform about a wider geographic area, as fundamentally different information will have been collected at each site. Interpretation of results in a local context may also be important. For example, a species that is nationally rare but locally common may indicate successful local management. Again, if only monitoring at one of these two scales, you would lose this information.

Similarly, sites defined for monitoring may be based on different contexts, such as ecologically relevant areas (e.g. river catchments), administratively relevant areas (e.g. counties or protected landscapes), areas related to potential funding streams (such as those covered by Wildlife Trusts) or those defined by relevant policies (e.g. LNRS areas), which also complicates potential to scale up in a consistent manner by combining data (see Appendix 1).

This could be resolved by ensuring local priorities are taken into consideration when targeting monitoring (see Section 3.2) and through using a system whereby local or regional coordinators feed in as part of a national governance structure to ensure there is enough flexibility for their individual priorities to be met (see Section 3.5).

2.5 Recorder motivation

Volunteers collect data for a variety of reasons, but these are often linked to their sense of place and their drive to contribute to local conservation priorities. These two motivating factors could be reduced if local projects were to begin feeding into multi-scale recording (Kühl *et al.* 2020). For example, feedback from volunteers in one project aiming to undertake landscape-scale monitoring has shown that a high level of connection to local strategic work is a strong motivating factor for volunteers (see Appendix 1). The need to use local volunteers to collect data at a national scale which matches the interests of the monitoring designers rather than the interests of the volunteers themselves has been described as a "tension" (Reichert *et al.* 2021). Other stakeholders involved, such as organisations that run local recording projects or academics undertaking local studies, may have similar reservations if they feel that they would not be able to publish themselves and would lose visibility if they were tied into a larger collaborative multi-scale monitoring initiative (Kühl *et al.* 2020). Recorder motivation can also lead to spatial biases in recording, for example increasing recording closer to volunteer homes and in areas with recording opportunities considered more interesting, such as areas with high species diversity (Dennis and Thomas, 2000). This can further exacerbate the problems relating to sampling design described in Section 2.1.

Such tensions could be reduced if data could effectively feed into both national and local datasets (see Section 3.5), and if volunteers are provided with feedback explaining how their data are being used (see Section 3.6). Augmenting volunteer recording with professional recording could also be considered in cases where priority at a larger scale is high but volunteer motivation is low (see Section 3.2).

3. Guidance to overcome these challenges

3.1 Understanding data requirements

When designing multi-scale monitoring, it is important to understand the data requirements that would be needed to support the potential applications of the data and ensure there is enough statistical power to answer relevant questions (Pescott *et al.* 2015). This will help to ensure that the data collected will be useable. It will also give an idea of how much recorder resource would be needed, which could help determine how realistic the design is likely to be in practice. As well as this, it can support other solutions such as targeting monitoring to where it will be most effective (see Section 3.2).

A number of studies have looked into data requirements for assessing trends at different scales. Having an average of at least 30 observations of a species each year is described as an “often used rule of thumb” to determine whether there is enough structured data to produce a trend (Border *et al.* 2019). One study (Pocock *et al.* 2019) developed “rules of thumb” to identify whether a set of occurrence records is likely to produce precise, and therefore useful, trends (noting that this only applies to occupancy-detection models). They concluded that “surprisingly few data” were required. To achieve 98% confidence, at least 29 records of the species in question were needed across the 10% best recorded years. Based on the data they were using (from UK recording schemes), this meant that on average only 12.3 records per species and year were needed. For species that were less commonly observed, this dropped to a minimum of ten records (averaging 4.5 per year). The study noted that the criteria used only provided evidence about the precision, not the accuracy or representativeness of the trends that were produced. It is also important to note that the amount of data required depends on how variable those data are (Larsen *et al.* 2001). However, it does suggest that the challenges highlighted above around an inability to draw smaller scale conclusions from datasets designed for a larger scale may not be as significant as many assume.

Another study (Henry *et al.* 2008) highlighted a similar need to understand data requirements. In this case, they were looking at how far apart sites could be to be able to use interpolation to predict biodiversity at sites in between the two. As biodiversity varies through time and space, they advised against extrapolation (i.e. applying information taken from one area to a larger area – also advised based on findings from Young *et al.* 2012) but promoted the idea of interpolation (i.e. if two of more sites are monitored, they can be used to infer information about sites located between them). This could be done, for example, by integrating data from two or more separate local schemes, to provide a larger scale trend. This would partly overcome the challenge explored in the previous section around an inability to draw regional or national conclusions from data collected for local monitoring purposes, in a way that does not use an unrealistic amount of extra recorder resource. However, understanding the limit to interpolation requires an analysis of the greatest distance at which there is spatial autocorrelation of the species presence or other characteristic being recorded. Where sites are located further apart than this distance, it will not be possible to use them in this way and further recording effort would be required. Carrying out this analysis to gain a good understanding of data requirements whilst designing a monitoring initiative would therefore be important. Spatial interpolation could also be used as an analytical technique to help integrate data across scales (Comber & Zeng 2019; Henry *et al.* 2008; Young *et al.* 2012), but is not repeated in Section 3.4 due to its inclusion here.

Other data requirements that would need to be considered when assessing data sources that could feed into multi-scale monitoring include how many years there are data for, how often sites are revisited, and how much variation there is across sites and species (Pescott *et al.* 2015). This could also lead to temporal autocorrelation which would also need to be analysed or considered.

It is important to note that understanding data requirements would be an iterative process (Stauffer *et al.* 2022). It is easier to assess how effective monitoring has been once the monitoring has begun producing data. Priorities may also change over time, which could lead to different data needs to those defined when the monitoring was initiated. Adapting the monitoring process based on updated understanding of data requirements going forwards would therefore be useful.

Understanding data requirements does not in itself break down the barriers described in the previous section, but it does allow us to better understand the extent of these barriers. In some cases, it may highlight the insignificance of barriers that were assumed to be there, and in other cases it may help to identify what is required to overcome the challenge.

3.2 Targeted monitoring

Targeting recorders to monitor in particular areas that are considered a priority across various scales could be used as a solution to ensure that data collected are as useful as possible (Callaghan *et al.* 2019). Currently, structured national schemes in the UK target recording towards a set of sites that have been designed to be representative of the wider country, through a random selection of sites that are representative of habitat type across the country. Local projects will also target monitoring to their area of interest depending on the question their monitoring is attempting to answer. However, if such a list of priorities could be combined and used to direct volunteers to high priority areas, this could create a dataset best able to answer all questions being asked. Whilst not all data would be useful to every question, the reasons a site was targeted could be used to pull out the best data from the set for any specific question. With the likelihood that at least some priorities would overlap, this would create a more efficient system, with one site visit potentially contributing towards multiple objectives, not just the one that the original recording scheme defined. This would require a good governance system (see Section 3.5) to oversee how priorities are set and to balance local and national interests.

As well as ensuring sites selected for national representativeness and sites selected due to their local interest (e.g. local conservation projects, areas earmarked for development) are covered, other factors that could be considered as priorities include ensuring an even spatial spread of data, temporal data gaps, and targeting sites that have had a previous visit to allow for temporal trends to be estimated (Border *et al.* 2019; Tulloch *et al.* 2013). In addition to targeting monitoring spatially as described above, taxonomic factors could be considered, such as functionally important taxa (e.g. pollinators), those sensitive to particular environmental pressures (e.g. indicator species), those typically under-recorded (e.g. fungi) or those of conservation priority (e.g. those classed as endangered on the IUCN Red List) (Border *et al.* 2019). For example, advice could be given that at site X, the highest priority taxa to record would be Y and Z.

An example of a project that is making use of modelling to target monitoring to areas considered high priority is DECIDE (Pocock *et al.* n.d.). This aims to produce fine resolution species distribution maps for a wide range of species (focusing on butterflies, moths, and grasshoppers to start with). The quality of the maps is continually improved by employing a process called adaptive sampling to nudge recorders to locations and at times that their recording will provide the most useful information to add certainty to the species maps. The exact adaptive sampling method can be tailored to not only improve the statistical models resulting from any collected data, but also to take volunteer interests and motivations into account. A related project, under JNCC’s Terrestrial Surveillance Development and Analysis project, is creating a 'Targeting Revisits Map' which uses a similar concept to encourage recorders to revisit sites in which only one year of data has currently been recorded, aiming to increase the utility of the data for trend analysis (Pocock *et al*. 2023).

Another example is the North American Bat Monitoring Program (NABat), which aims to address the "tension" between top-down and bottom-up monitoring in a number of ways, including through a "hierarchical master sample survey design" (Reichert *et al.* 2021). They have developed a web-based mapping tool that directs users to record at the highest priority grid cells within the region they select, based on a set sampling order as defined by the NABat Master Sample. The Master Sample uses a similar sampling design to the UK's national recording schemes that were described previously. Sites are selected to contain a representative amount of different habitat and land use types but are selected randomly within each of these types (stratified random sampling). However, instead of selecting a set number of squares, it determines a priority order for squares. Thus, the more recorders participate, the smaller the scale the data can be applied to and analysed over. The similarities to the current UK data landscape suggest potential for relatively minor adjustments to the current system that could lead to greater potential for multi-scale application. Local scale studies can also be nested within the NABat Master Sample grid. For example, a local scale project in Crater Lake National Park that aimed to investigate the response of bats to their forest wildfire fuel reduction programme was allocated extra grid squares. Some of these were squares needed for the national analysis and others provided extra data only for the local project (Reichert *et al.* 2021).

One study (Wright *et al.* 2022) assessed the effectiveness of different monitoring designs to estimate trends in wildlife communities across multiple spatial scales. Stratified random sampling was found to be the most useful of the designs tested when being considered in the context of its ability to provide information across spatial scales, regardless of sampling effort. Stratified random sampling should therefore form a strong component of the priorities set for targeted monitoring in a multi-scale framework. However, the study also identified that other sampling designs performed better under specific situations (for example, the reasons behind regional level changes were found to be best predicted when using a different 'rotating panel' sample design), giving support to the idea of needing to leave flexibility in the system for more specific local needs to be met.

In some cases, the reason an area or a taxon is less well recorded than others may be to do with technical barriers to monitoring. Combining targeted monitoring with novel monitoring techniques, which is becoming increasingly possible with recent advances, could be useful to solve this problem (Border *et al.* 2019). For example, sampling environmental DNA could help identify the presence of taxa that are too small to see or are difficult to identify visually (Thomsen & Willerslev 2015). Similarly, Earth Observation data (e.g. satellite imagery) are useful to assist with the creation of habitat maps (Sparrow *et al.* 2020), and the use of passive recorders (e.g. acoustic monitoring, camera traps) help to record species that are easily disturbed by human presence, elusive, or normally active at night (Kelly 2008; Sugai *et al.* 2019).

In other cases, data gaps may have arisen from areas that are remote or less convenient for volunteers to travel to. Alternatively, areas with apparent data gaps may simply have fewer of the target species, which is just as useful to record from a statistical perspective but may seem less motivating to the individual carrying out the recording (Pescott *et al.* 2015; Reichert *et al.* 2021). In such cases, it may be useful to augment volunteer recording with professional recording (Border *et al.* 2019; Reichert *et al.* 2021).

In summary, if a set of priorities could be agreed upon and weighted (in terms of where and what should be targeted to monitor), this could provide a strong contribution to overcome several of the challenges that multi-scale monitoring faces, through providing a sampling design that is relevant across different scales, an ability to take into account differing local priorities, and an ability to answer multiple questions from the same dataset.

3.3 Use of analytical techniques

Whilst standardising protocols across initiatives would be the simplest solution to monitoring across scales, it is unlikely to be feasible to implement in the real world. Therefore, making best use of data collected through different protocols using analytical techniques will also be important, which advances in statistical modelling are increasingly able to facilitate (e.g. Mancini *et al.* 2022; Zipkin & Saunders 2018).

Many of the challenges outlined in Section 2 were reflective of limitations to traditional analytical techniques and statistical practices. For example, the inapplicability of sampling design across scales (Section 2.1) and the inconsistency of protocols between different monitoring initiatives (Section 2.2) are problems because of the requirements of unbiased and comparable data in traditional statistical analysis. However, a range of alternative analytical techniques can be implemented that could solve such problems. These include reducing data to their smallest common denominator, calibration, inverse stratification, integrated modelling, meta-analysis, and interpolation of missing data (Henry *et al.* 2008; Kühl *et al.* 2020; Mancini *et al.* 2022; Pescott *et al.* 2015; Young *et al.* 2012), each of which are explored below. While solutions such as targeted monitoring (Section 3.2) and encouraging interoperable protocols (Section 3.4) aim to improve the data landscape in such a way that traditional analyses are possible, the solution described here focuses on making best use of an imperfect data landscape (Border *et al.* 2019). These two types of solution can be used in combination to address the problem from both directions at once.

The simplest option to allow for analysis to take place despite inconsistent protocols is to simply reduce data to their smallest common denominator. For example, if one monitoring initiative recorded presence/absence of a particular species and another initiative recorded counts detailing how many of that species were present at each observed location, data on presence/absence could be extracted from the second one to give a comparable dataset that could be combined with the first (Henry *et al.* 2008). A key disadvantage of this approach is that information is lost from the more detailed initiative, and effort thereby ‘wasted’, which might be demotivating for volunteers. It also leads to a risk of false negatives being inferred in cases where recording absence is not a key aim (Adams & Muths 2019).

Another analytical option that can be used in some cases to combine datasets that initially seem incomparable is calibration (Henry *et al.* 2008). Calibration, defined as correcting for the variation in the information obtained from two different data collection processes, requires a part of the two datasets to overlap, or targeted monitoring to add some overlapping data. For example, if two different methods are used to monitor the population of a particular species, undertaking both methods at a small number of the total sites could lead to an understanding that one method consistently overestimates compared to the other by X%. This information could be used to ‘correct’ the data in both datasets in a way that reduces estimates from one and increases estimates from the other, to try to give more comparable results. However, it is never possible to be completely certain that the calibration has adjusted ‘correctly’. An example of where a concept like this has been applied was when moving from the Common Bird Census methodology to the current BBS methodology; both were run for seven years and checked to ensure that they gave the same year-to-year pattern in the data (Freeman *et al.* 2007).

Inverse stratification is an analytical technique that allows for an uneven distribution of habitat types to be present in the data without biasing the results (Henry *et al.* 2008). For example, if one type of habitat is sampled more than another, it can be given a lower weighting in the analysis of results to ensure that the final estimate is representative. This is automatically accounted for in many model types, such as GLMs (Generalised Linear Models). This would resolve the challenge identified in Section 2.1 around sampling design. However, if methods differed across points sampled (Section 2.2) this would remain an issue.

Model-based data integration (or integrated distribution modelling when applied to species distributions) provides another option to combine data from multiple sources (Mancini *et al.* 2022). This is a technique that analyses data collected using different protocols jointly but uses a separate observation model for each. Observation models describe the generation of species records conditional on both the spatial distribution of the species, and the spatial distribution of the recording effort and other related factors such as seasonality This makes it possible to integrate the two data sources to estimate species occupancy, abundance and/or trends, in a way that can produce more precise estimates and correct for biases in the observation process (Adams & Muths 2019; Kühl *et al.* 2020). However, it remains a computationally intensive method that requires technical expertise to undertake (Mancini *et al.* 2022). Model-based data integration is particularly suited to combining data from structured and unstructured monitoring initiatives, which is useful as it allows for the benefits of both data types (namely the abundance and ease of collecting data in unstructured monitoring, and the statistical rigour and extra information provided by structured monitoring) to be retained (Border *et al.* 2019; Mancini *et al.* 2022). However, an appropriate balance of data between the two are required, and when the different data sources provide contrasting information (e.g. because different areas or habitats have been sampled), the model may struggle to reconcile them, which may result in a less precise estimate (Jarvis *et al.* 2021; Mancini *et al.* 2022). This highlights the need for better metrics to measure model performance besides precision. Integrated models are typically implemented using a Bayesian (probability-based) statistical framework, as this provides a flexible way to combine data collected at different scales and through different methods (Henry *et al.* 2008; Mancini *et al.* 2022; Pescott *et al.* 2015).

Meta-analyses are another form of analytical technique that can be used to combine different forms of data (Henry *et al.* 2008). In these cases, the outputs from independent monitoring are used as inputs to analyse a more general pattern. Meta-analysis is typically used where there is interest in matching an observed variable with an independent variable to understand correlation, and they are a common approach for creating literature reviews to answer this type of question. For example, if a set of studies had independently measured the population of a particular species and recorded a factor related to climate change such as temperature at each site, meta-analysis could be used to draw conclusions from the two datasets about the average effects of climate change on this species' population. This would be possible whether the original methods and sampling were the same or not (Henry *et al.* 2008). However, this requires studies to adequately document their methods (so the impact of any differences may be understood) and clearly state effect sizes and associated errors, which monitoring outputs should include, whatever their scale.

In summary, the use of analytical techniques to integrate datasets collected using different methods and sampling approaches can make best use of the current situation, pulling on the strengths and complementarity of a wide array of different monitoring initiatives (Sparrow *et al.* 2020).

3.4 Encouraging interoperable protocols

Protocols can be considered interoperable when it is possible to use the data that they produce in an integrated way. As explored in the previous section, it is not necessary for protocols to be identical and fully standardised for this purpose, as analytical approaches provide options for combining different types of data. However, a certain level of similarity is required, particularly in terms of coding of covariates, to be able to combine data from different protocols in this way. Encouraging monitoring initiatives to adopt interoperable protocols is therefore important if using their data for application across multiple scales.

A key theme emerging from the literature is that an effective way to encourage interoperable protocols is to create a framework of national guidelines or standards, which allow for local flexibility (Adams & Muths 2019; Danielsen *et al.* 2005; Kühl *et al.* 2020; Thomaes *et al.* 2021). The specific methods used at a local scale can therefore be determined by individual initiatives if they fit within the guidelines. This would require strong collaboration and coordination between local and national actors (see Section 3.5). It is important for all stakeholders to be involved in the process of co-designing the overarching rules and norms (see Section 3.5), and to agree on the degree of local flexibility they will require for each of their use cases. This is necessary to foster a culture where such integration is likely to be sustainable in the long term (Kühl *et al.* 2020).

One example of an initiative that has taken this approach is NABat (introduced in Section 3.2). NABat expects partners to follow national guidance, but also has regionally specific standard operating procedures (Reichert *et al.* 2021). Another example is the European Stag Beetle Monitoring Network (ESBMN), which created an internationally standardised protocol that allowed for some local variation (Thomaes *et al.* 2021). Some local variation was required because the phenology of the species (and therefore the optimal time of year to sample) varies throughout its range. Other variation arose whilst meeting the requirements and priorities of local actors.

Alternatively, ensuring that a small number of methods (each of which are well replicated) are defined and undertaken could provide another route for ensuring effective interoperable protocols (Danielsen *et al.* 2005). For example, if a set of standardised protocols are developed from which local initiatives can select an option to follow, this leaves local initiatives with the flexibility to select the one that is most relevant to their local question. In doing so, this contributes to larger scale questions as well, making it easy to combine their findings with those from other similar studies taking place elsewhere. This is a particularly useful approach when applying meta-analytical techniques (Danielsen *et al.* 2005). Again, it would be important to involve stakeholders with an interest in each of the different scales when deciding upon the set of protocols to be used (see Section 3.5).

Encouraging interoperability can also take a lighter touch approach. As long as robust and unbiased estimates of common parameters or attributes (e.g. occupancy, abundance, survival) are being made across different initiatives, this can be enough for integrated analyses to take place (Adams & Muths 2019; Danielsen *et al.* 2005). One example of a monitoring initiative that has taken this approach is the Amphibian Research and Monitoring Initiative (ARMI). In setting up the project, the scheme's designers did not want to impose a standardised methodology that would limit their partner's abilities to answer management questions (Adams & Muths 2019). Therefore, no limitations, guidance or criteria were set, apart from setting robust estimates for a group of common parameters that should be measured. Local initiatives could therefore choose where to carry out monitoring and whether to carry out their monitoring through trapping, visual or other methods, depending on what was most useful or relevant to the local question being asked in their study. However, all this information was found to be useful at a later stage when ARMI combined data on each parameter from across the studies undertaken to give information at a larger scale, even where different methods were used to collect information on the same parameter, with no additional analysis required. Whilst early ARMI activity focused on monitoring tailored to address local research questions, they were later able to start synthesising information from across the programme to understand larger scale trends (which was always the ultimate ambition). They recommend forming a common framework for addressing questions and aligning terminology used across initiatives at the outset, in order to facilitate the subsequent scaling up of findings (Adams & Muths 2019).

At a UK scale, a similar approach has been taken by the four Governments of the UK within Common Standards Monitoring (JNCC, n.d.). This was set up to provide an agreed way to review the condition of statutory sites across England, Wales, Scotland, and Northern Ireland. The framework provides guidance around the assessment of designated interest features, each of which is based on gathering data against defined attributes which describe the features condition. A flexible approach is adopted for the monitoring methods used, with monitoring data assessed against the defined attributes to rank features into the categories favourable, unfavourable, or partially/fully destroyed. The use of categories rather than raw data means there is a common framework for comparison across all features and countries that could be flexible to the precise monitoring methods used. These protocols therefore allow for interoperability without restricting those undertaking local monitoring to a prescribed protocol.

Another approach suggested to increase the interoperability of protocols involves standardising the methodology but not the sampling approach. An example of this is the 'Tracking the Impact' study taking place in the Chilterns Area of Outstanding Natural Beauty in the UK, which has replicated the data collection protocols of three national taxa schemes (the Breeding Bird Survey, the Wider Countryside Butterfly Survey, and the National Plant Monitoring Scheme). The sampling approach that they are using is designed to give landscape-scale trends rather than national trends, thus while some of the survey squares overlap with the national schemes, many others have been added within the area of interest to allow for a smaller scale analysis (see Appendix 1 for further information). Whilst not contributing to the national scale trends as it would bias the national sampling protocol geographically, the fact that the same methods are used could make comparisons and integration using analytical techniques possible in the future. Approaches such as this that are based on standardising methodology lend themselves more to new monitoring initiatives that are being set up, rather than to making best use of data that is already collected; this avoids disturbing the initial aims of the monitoring, confusion among recorders, and comparability across time series (Stauffer *et al.* 2022).

Encouraging interoperable protocols is essential to avoid problems described in Section 2.2 around incomparability but must be done so in a way that allows for specific questions (see Section 2.3) and local contexts (see Section 2.4) to also be considered. A flexible framework that aligns key aspects but relies on analytical techniques to avoid the need for complete standardisation is therefore likely to be the best approach.

3.5 Ensuring complementarity, collaboration, and coordination across scales

Monitoring across scales provides significant logistical challenges. There is therefore a need for coordination and collaboration across the different scales and organisations that would be involved. A clear governance structure would be required to ensure that everyone had a platform to discuss their needs and agree on how best to meet the requirements across different groups, and to ensure that everyone understood their responsibilities within the overall operation. Such a system would avoid dividing or diluting recording effort across different initiatives, so that as much data as possible is available centrally for as many applications as possible.

A number of studies suggested a governance structure consisting of a national body, with local and regional coordinators (Kühl *et al.* 2020, 2020; Pavlacky Jr *et al.* 2017; Reichert *et al.* 2021; Tulloch *et al.* 2013). For example, NABat has a 'Core Team' made up of representatives from a number of national government departments, who were responsible for programme coordination and support, data management and analysis, development of statistics, and IT (Reichert *et al.* 2021). It also has an increasingly large set of regional monitoring hubs that coordinate monitoring efforts on the ground. The actors involved in these vary, but include conservation agencies, tribal land management agencies, universities and NGOs (Reichert *et al.* 2021). Similarly, the Scottish Biodiversity Information Forum (SBIF), which was established in 2012 to reduce challenges relating to data mobilisation in Scotland, recommended the introduction of a governance structure with regional and national hubs (Kühl *et al.* 2020). Another example of such an approach, (although slightly different in context being a legal framework rather than a voluntary monitoring system), is the EU's Habitat Directive. The Directive provides an overarching framework for monitoring in which to take place, but is based on an integration of reporting that is already taking place in each country (Henry *et al.* 2008).

The high-level national or international body was seen as most effective for implementing tasks such as providing direction, coordinating the development of a framework against which the monitoring can take place (Reichert *et al.* 2021), data management and analysis (Thomaes *et al.* 2021) and targeting monitoring (Henry *et al.* 2008). Coordination to overcome other barriers, such as minimising effort by ensuring that landowners are only approached for access requests once with a request covering all monitoring activity within the framework (rather than repeated separate requests from different monitoring initiatives) could also be best undertaken by the high-level body.

Meanwhile, the local and regional coordinators would be most efficient at tasks such as recruiting, training and coordinating local volunteers (Thomaes *et al.* 2021), ensuring data quality of records submitted by volunteers (Tulloch *et al.* 2013), and adapting monitoring protocols to ensure that they are flexible enough to meet the requirements of local stakeholders (Reichert *et al.* 2021). Volunteers feel most associated with the local level of monitoring, so any volunteer interaction tasks are best carried out by local coordinators (Thomaes *et al.* 2021).

Both the high-level body and the local or regional coordinators would need to be involved in data flows to ensure that the data collected reach the place they can be used. This can be challenging so would need design of good infrastructure and good communication throughout all stages of the process.

Key to the success of both the design and implementation of multi-scale monitoring is the inclusion and input of a wide variety of stakeholders (Sparrow *et al.* 2020). Insight would be required on the needs of each scale, the local context of each organisation that would be involved, and the needs of different types of actors likely to be involved (NGOs, volunteers, analysts, environmental consultants, local naturalist groups, landowners, etc.). Each of these groups is also likely to bring a different skillset that they can contribute to the monitoring design (Sparrow *et al.* 2020). Inclusion such as this would ensure that different monitoring initiatives can be brought together in a way that makes them greater than the sum of their parts, rather than losing the needs of certain actors in an effort to centralise things (Kühl *et al.* 2020). For example, the NABat project consults a range of actors through regional and technical working groups, to help make improvements to the recording system (Reichert *et al.* 2021). Similarly, the Integrated Monitoring in Bird Conservation Regions partnership is made up of representatives from national and regional government, NGOs, academia and Native American Nations, thus taking in a wide range of views during the design process (Pavlacky Jr *et al.* 2017).

Overall, a strong and representative governance structure and willingness to collaborate are likely key to successful design and implementation of multi-scale monitoring.

3.6 Feedback to volunteers

Providing feedback to volunteers regarding how the data they are collecting is being used, is important when monitoring at any scale to ensure continued volunteer motivation. However, when implementing a solution that involves integrating data from current monitoring initiatives to provide information for a different purpose to that for which the initiative was originally designed (such as at a different scale), it is likely to hold even more importance. For many volunteers, the motivation to collect data is associated with an attachment to their local area (see Appendix 1). If it were perceived that the data they were collecting was not being used as effectively for its original purpose because of compromises that had to be made to fit into the overarching framework, or if volunteers lost the sense of place that their monitoring is associated with because the data collection was presented as too centralised, this could lead to reduced levels of volunteer motivation, and subsequently reduced participation. Communicating the utility of the data collected could be undertaken, for example, through newsletters or a dedicated online site (Tulloch *et al.* 2013). This could include data about the new additional uses to which the data are being put but should also reassure volunteers that their original motivation for taking up recording is still being met. This aspect is something that NABat and ESBMN have recognised as a potential future improvement to their multi-scale monitoring processes (Reichert *et al.* 2021; Thomaes *et al.* 2021).

4. Implementation advice

4.1 Advice for those designing multi-scale recording

1. **Define the overarching scope**

    - Which taxon/taxa do you plan to cover?
    - Do you plan to cover all scales (local to global), or a defined subset of scales (e.g. regional and national)?
    - Do you have any other key boundaries from a top-down perspective that will narrow the set of stakeholders you will be engaging with? (e.g. perhaps you have a key interest in farmland rather than all land use types, or your scope will be geographically restricted to a certain country or region).

2. **Identify relevant stakeholders**

    - Potential data providers, for example:
        - Current recording initiatives that overlap in scope (those organising the scheme and those participating in the scheme).
        - Local interest groups.
        - Potential citizen scientists (e.g. through social media).
        - Professional recorders.
        - Local Environmental Records Centres.
    - Potential data users, for example:
        - Those implementing local conservation projects.
        - Those participating in agri-environment schemes.
        - Policymakers at a variety of scales (local government, national government).
        - NGOs.
        - Academics.

3. **Understand stakeholders' needs and priorities**

    - Set up workshops, surveys, or interviews to understand the motivations of potential data recorders and the kinds of questions that potential data users hope to answer.
    - Ensure this is an iterative process of codesign and there is a mechanism for continued input about how well the monitoring is meeting these needs and priorities once implementation has started, and as new needs and priorities arise.

4. **Set up a governance system**

    - This should include both top-down representation (the organisation setting up the overall monitoring) and bottom-up representation (the stakeholder groups from the previous section), to integrate a wide range of priorities into the decision-making process.

- The bottom-up group should have good geographical representation that can form a network of local coordinators.
- The roles of each member of the governance structure should be defined. The top-down representatives are likely to be most effective coordinating aspects such as data infrastructure, and the bottom-up representatives should focus on local engagement.
- Strong communication channels should be established, such as a regular meeting for discussion.

5. Finalise the scope and plan the implementation phase

- What questions does the recording aim to answer? This is likely to be a list of questions specified by and agreed between multiple stakeholders.
- How much data will be required to answer these questions? Consider rules of thumb (see Section 3.1) and the size of the areas data will be required from to answer the questions specified.
- What analysis will be required to answer these questions (see Section 3.3 and step 11 in this list)?

6. Set out acceptable protocol(s)

- If rigour, consistency, and comparability are the key priorities, this could be a single protocol that is replicated everywhere.
- If flexibility, use by the widest number of local use cases possible and engagement of the greatest number of volunteers is the priority, this could include any method for recording a common variable.
- If the priority is a combination of the two, this could be a small set of defined protocols. Each analysis could choose whether to use data from across the set of protocols or from just one.
- Protocols could include optional data collection. For example, if one of the priorities is to establish whether there is a correlation between management option X and the biological data being recorded, but a range of other questions are also being answered, then the option to record whether management option X is taking place at any given sampling point could be included in the protocol. Only those that do record this information would be included in the relevant analysis for this question, but all samples could be included when answer some of the more general questions.
- Many protocols already exist. It is likely to be beneficial to align with these as closely as possible (unless this is not possible due to the type of question being asked) as the monitoring could then pull data from existing collection routes rather than 'reinventing the wheel'.

7. Define a sampling strategy

- If statistical rigour is the priority (e.g. required for an official statistic) AND only larger scales are of interest (e.g. comparing regional and national trends), the sampling strategy should be based on stratified random sampling. It may not be feasible to use this sampling strategy if multiple and very different scales are within scope (e.g. national coverage, and a selection of local sites), as the density of recording that would be required to cover the larger scale to the same sampling strategy of the smaller scales would be very high.

- If engagement of the greatest number of volunteers is the priority, an ad hoc sampling approach could be taken (they can record anywhere). This could be corrected using inverse stratification or combined with other types of recording through integrated modelling (see Section 3.3).
- If a range of priorities have been defined, including answering more local-scale questions, a targeted monitoring approach could be taken that directs people to areas of interest. This could include a set of random stratified samples for use in analyses requiring the highest statistical rigour and a set of samples identified due to specific local priorities (e.g. areas undergoing conservation actions). Each analysis could choose whether to use data from across all data collected or just data collected from samples prioritised for a certain reason (e.g. national trend analysis could only use data from the random stratified set).

8. Identify data collection routes

Depending on the questions being asked:

- It may be possible to combine data already being collected by current recording initiatives.
- It may be possible to assist current recording initiatives to expand or adjust collection routes to fit into the multi-scale monitoring.
- It may be necessary to recruit new volunteers and set up new systems for data collection.
- It may be useful to augment volunteer recording with professional recording if funding is available and some of the priority areas are remote or unmotivating to volunteers.

9. Set up the data infrastructure

- A central storage system (or system that is able to access and pull from a range of decentralised storage questions) would be required. Ideally this would be open access, allowing all current and potential future stakeholders to make full use of all possible opportunities to use the data.
- Data flows that allow efficient transfer of information from the point of collection to the central storage system would be required.
- The data infrastructure would need to support all information defined in the previous sections. For example, if targeted monitoring for a range of priorities is taking place, each sample would need to be tagged with the reason for their priority so that those analysing the data can make their own decisions about which parts of the dataset to use. Similarly, if a priority question is to answer whether there is a correlation between management option X and the biological data being recorded, the ability to filter samples that have recorded whether management option X has taken place at the site would be required.
- Data validation will be an important step in the process. If aligning with existing data collection processes, this may already be in place. If not, it will need to be set up.
- Data management, GDPR, data storage costs and security implications will need to be considered.

- As well as storing the raw data, presentation of data post analysis (e.g. through an interactive dashboard) could be considered as it would be useful to communicate data uses and findings back to recorders for motivation or to those using the data to inform decision making.

10. Encourage participation

- This would best be done by local coordinators.
- Stakeholders identified in step 2 of the list could be engaged.
- Promotion on social media or at local events could be useful.
- Training and support (e.g. local WhatsApp groups) could be provided.
- Clear communication of what the data will be used for and how valuable it is will be important. This could be adapted to the audience.
- Continued and clear communication of what the data has been used for, and how valuable it has been, will be important to ensure retention of volunteers. They are likely to be especially motivated by their contributions to local questions and to solving local problems, so this feedback could be tailored by local coordinators.

11. Analyse the data

- For each of the questions defined in step five of this list, undertake appropriate analysis. This will vary depending on the protocols and sampling strategies set out in steps 6 and 7.
- For monitoring in which a small set of different protocols are set out and a combination of sampling strategies are used, integrated modelling is likely to be a strong candidate to enable robust analysis of the overall dataset, if appropriate computational power and analytical expertise are available.
- For monitoring in which any protocol measuring the same parameter is permitted and questions typically focus on identifying correlation between records of a species and a particular factor such as a management action or climate change, meta-analysis may be the most appropriate analytical technique to use.
- Some questions may require a simpler and more standard analysis of just part of the dataset, e.g. where relevant information has been recorded optionally or where only data from a particular geographic area is relevant.
- Local analysis should be possible to contextualise within the national picture to provide greater insight and relevance.
- More detail on analytical options is outlined in Section 3.3.

12. Obtain feedback on data uses

- Consult stakeholders both within and outside of the governance structure to understand whether the data being collected are as useful as it could be. This could be done through surveys or workshops.
- Where feedback suggests feasible improvements to the monitoring, these should be implemented going forwards.

13. Iterate all steps to ensure the recording framework continually improves, expands, and remains relevant to future use cases as they arise

4.2 Advice for those running existing recording initiatives

1. **Consider the scale(s) at which your current data recording is relevant.** Perhaps your recording is targeted at one scale, but by using different types of analyses could provide relevant information to another scale.
2. **Consider the scales at which your current data recording could contribute to.** Perhaps your recording is restricted to a particular geographic region, but if collaborating with those recording in other geographic regions could contribute towards a larger scale trend. Perhaps your recording aims to create national trends through data collection from a set of randomly stratified sites, but some of these sites overlap with SSSIs and so could be of interest to those monitoring or managing these areas.
3. **Identify other organisations collecting similar data and consider whether there would be value-add to combining the data you collect.** For example, two neighbouring interest groups could combine the data they collect to help calculate regional trends in that area and compare data between the two. Similarly, a national scheme recording a particular taxon could link up with local projects doing the same to create a larger database that allows analysis of a greater range of questions. If an organisation hoping to develop a multi-scale monitoring exists, following a similar process to that described in Section 4.1, you could feed into the stakeholder engagement and governance structure.
4. **Consider whether additional data could be added to your protocols**. For example, if an organisation you are considering combining data with records slightly different parameters, such as habitat in addition to species counts, you could consider adding a question in your protocol (even if this is optional to avoid additional burden) to support additional data to the questions they are trying to answer for minimal extra effort in your current scheme – and spot similar opportunities that could help collect more data for your own questions. See also Broughton and Pocock (2022).
5. **Communicate multi-scale uses of the data**. Volunteers are motivated by hearing how useful their data has been. Adding variety to the applications that their data is used for is therefore likely to add to motivation. While volunteers are typically more motivated by local applications, hearing about multi-scale uses (both local and national) would ensure that they learn about and appreciate higher level data needs whilst reassuring that they continue to contribute just as much to more local data needs.

4.3 Advice for those designing monitoring for local or regional conservation projects

1. **Define the question that is being monitored against and understand data requirements**. See advice in Sections 4.1 and 3.1.
2. **Identify and engage with relevant stakeholders**. This is likely to be a narrower set than those required in Section 4.1, as the scale has already been defined. However, understanding the motivations of those who may be collecting the data and the requirements of those likely to use the data will be key. Inclusion of representatives outside of the directly intended users, for example who may be interested in using the data at a national scale, would be useful to include. Whilst priority should be given to the key users, if insight from these ‘secondary’ users can lead to changes in the design that have minimal effect on the effort required to set up or undertake the

monitoring and significant benefit to potential data users at a larger scale, their requirements should also be considered.

3. **Review monitoring protocols and sampling strategies that have been undertaken in similar conservation projects previously**. Where appropriate, selecting one of these rather than designing your own will both save time, and improve potential for combining data.
4. **Contextualise results from your project with regional or national data**. Analysing data at multiple scales, either through aligning protocols or through implementing one of the analysis techniques described in Section 3.3, will help to interpret results from the monitoring of your local project. For example, an increasing trend in your project area may be more significant if the same factor is decreasing nationally. Alternatively, if your local project is showing an increase in a particular species, it would be worth investigating whether the conservation actions being undertaken at your site are truly increasing populations or whether they are simply attracting individuals from the nearby landscape.

5. Conclusions

In conclusion, whilst there are substantial challenges to multi-scale monitoring and to combining data from multiple recording initiatives to achieve common aims, it is also an area that shows significant potential for improving environmental recording, and for increasing the applicability and efficiency of data collected. Increasing collaboration and data sharing between existing initiatives, ensuring that new initiatives are designed with alignment and shared analysis in mind, using analytical techniques that help correct for differences in data collected by different initiatives, and establishing multi-scale monitoring that brings together stakeholders and creates data infrastructure that can be contributed to and used by those recording at all scales will all help break down current barriers to multi-scale monitoring and help inform conservation across multiple scales in the future.

References

Adams, M.J. & Muths, E. 2019. Conservation research across scales in a national program: How to be relevant to local management yet general at the same time. *Biological Conservation* 236, 100–106. https://doi.org/10.1016/j.biocon.2019.05.027

Ascensão, F., Barton, M., Crossman, K., Edgecomb, M., Elmqvist, T., Gonzalez, A., Guneralp, B., Haase, D., Hillel, O., Huang, K., Maddox, D., Mansur, A., Paque, J., Miguel Pereira, H., Rae Pierce, J., Weller, R., Seto, K., Mei Jia Tan, M. & Ziter, C. 2018. Nature in the Urban Century. The Nature Conservancy.

Border, J., Gillings, S., Newson, S., Logie, M., August, T., Robinson, R. & Pocock, M. 2019. The JNCC Terrestrial Biodiversity Surveillance Schemes: An Assessment of Coverage. JNCC Report No. 646, JNCC, Peterborough, ISSN 0963-8091. [WWW Document]. URL https://hub.jncc.gov.uk/assets/c3b082a9-7e9e-4e8e-ae1e-fd80e1bdbbab (accessed 1.9.23).

Broughton, R. & Pocock, M. 2022. The opportunities for semi-structured and effort recording to enhance the value of biological recording by volunteers. JNCC, Peterborough. https://hub.jncc.gov.uk/assets/57e64ede-78b2-43dd-8e50-35555874fdbc

Callaghan, C.T., Rowley, J.J.L., Cornwell, W.K., Poore, A.G.B. & Major, R.E. 2019. Improving big citizen science data: Moving beyond haphazard sampling. *PLOS Biology* 17, e3000357. https://doi.org/10.1371/journal.pbio.3000357

Comber, A. & Zeng, W. 2019. Spatial interpolation using areal features: A review of methods and opportunities using new forms of data with coded illustrations. *Geography Compass* 13, e12465. https://doi.org/10.1111/gec3.12465

Cowie, R.H., Bouchet, P. & Fontaine, B. 2022. The Sixth Mass Extinction: fact, fiction or speculation? *Biological Reviews* 97, 640–663. https://doi.org/10.1111/brv.12816

Danielsen, F., Burgess, N.D. & Balmford, A. 2005. Monitoring Matters: Examining the Potential of Locally-based Approaches. *Biodivers Conserv* 14, 2507–2542. https://doi.org/10.1007/s10531-005-8375-0

Dennis, R.L.H. & Thomas, C.D. 2000. Bias in Butterfly Distribution Maps: The Influence of Hot Spots and Recorder's Home Range. *Journal of Insect Conservation* 4, 73–77. https://doi.org/10.1023/A:1009690919835

Freeman, S.N., Noble, D.G., Newson, S.E. & Baillie, S.R. (2007) Modelling population changes using data from different surveys: the Common Birds Census and the Breeding Bird Survey, *Bird Study*, 54:1, 61-72, DOI: 10.1080/00063650709461457

Henry, P.-Y., Lengyel, S., Nowicki, P., Julliard, R., Clobert, J., Čelik, T., Gruber, B., Schmeller, D.S., Babij, V. & Henle, K. 2008. Integrating ongoing biodiversity monitoring: potential benefits and methods. *Biodivers Conserv* 17, 3357–3382. https://doi.org/10.1007/s10531-008-9417-1

IPBES. 2019. Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services. IPBES Secretariat.

Jarvis, S.G., Staley, J.T., Siriwardena, G.M., Seaton, F., Redhead, J.W., Ward, C., Upcott, E., Botham, M.S. & Henrys, P. 2021. Modelling Landscape-scale Species Response to AgriEnvironment Schemes.

JNCC (n.d.) Common Standards Monitoring | JNCC - Adviser to Government on Nature Conservation [WWW Document]. URL https://jncc.gov.uk/our-work/common-standards-monitoring/ (accessed 2.21.23).

Kelly, M.J., 2008. Design, evaluate, refine: camera trap studies for elusive species. *Animal Conservation* 11, 182–184. https://doi.org/10.1111/j.1469-1795.2008.00179.x

Kühl, H.S., Bowler, D.E., Bösch, L., Bruelheide, H., Dauber, J., Eichenberg, David., Eisenhauer, N., Fernández, N., Guerra, C.A., Henle, K., Herbinger, I., Isaac, N.J.B., Jansen, F., König-Ries, B., Kühn, I., Nilsen, E.B., Pe'er, G., Richter, A., Schulte, R., Settele, J., van Dam, N.M., Voigt, M., Wägele, W.J., Wirth, C. & Bonn, A. 2020. Effective Biodiversity Monitoring Needs a Culture of Integration. *One Earth* 3, 462–474. https://doi.org/10.1016/j.oneear.2020.09.010

Larsen, D.P., Kincaid, T.M., Jacobs, S.E. & Urquhart, N.S. 2001. Designs for Evaluating Local and Regional Scale Trends: We describe a framework for evaluating the effects of spatial and temporal variation on the sensitivity of alternative ecological survey designs to detect regional temporal trends. *BioScience* 51, 1069–1078. https://doi.org/10.1641/0006-3568(2001)051[1069:DFELAR]2.0.CO;2

Mancini, F., Boersch-Supan, P., Robinson, R., Harris, M. & Pocock, M. 2022. An introduction to model-based data integration for biodiversity assessments. JNCC, Peterborough, ISBN 978-1-86107-639-7. [WWW Document]. URL https://hub.jncc.gov.uk/assets/1c774649-3cf8-4964-bf38-443a12accd09 (accessed 1.9.23).

Navarro, L.M., Fernández, N., Guerra, C., Guralnick, R., Kissling, W.D., Londoño, M.C., Muller-Karger, F., Turak, E., Balvanera, P., Costello, M.J., Delavaud, A., El Serafy, G., Ferrier, S., Geijzendorffer, I., Geller, G.N., Jetz, W., Kim, E.-S., Kim, H., Martin, C.S., McGeoch, M.A., Mwampamba, T.H., Nel, J.L., Nicholson, E., Pettorelli, N., Schaepman, M.E., Skidmore, A., Sousa Pinto, I., Vergara, S., Vihervaara, P., Xu, H., Yahara, T., Gill, M. & Pereira, H.M. 2017. Monitoring biodiversity change through effective global coordination. *Current Opinion in Environmental Sustainability* 29, 158–169. https://doi.org/10.1016/j.cosust.2018.02.005

Niemelä, J. 2000. Biodiversity monitoring for decision-making. *Annales Zoologici Fennici* 37, 307–317.

Paloniemi, R., Apostolopoulou, E., Primmer, E., Grodzinska-Jurcak, M., Henle, K., Ring, I., Kettunen, M., Tzanopoulos, J., Potts, S., Hove, S. van den, Marty, P., McConville, A. & Simila, J. 2012. Biodiversity conservation across scales: lessons from a science–policy dialogue. *Nature Conservation* 2, 7–19. https://doi.org/10.3897/natureconservation.2.3144

Pavlacky Jr, D.C., Lukacs, P.M., Blakesley, J.A., Skorkowsky, R.C., Klute, D.S., Hahn, B.A., Dreitz, V.J., George, T.L. & Hanni, D.J. 2017. A statistically rigorous sampling design to integrate avian monitoring and management within Bird Conservation Regions. PLOS ONE 12, e0185924. https://doi.org/10.1371/journal.pone.0185924

Pescott, O.L., Walker, K.J., Pocock, M.J.O., Jitlal, M., Outhwaite, C.L., Cheffings, C.M., Harris, F. & Roy, D.B. 2015. Ecological monitoring with citizen science: the design and implementation of schemes for recording plants in Britain and Ireland. *Biological Journal of the Linnean Society* 115, 505–521. https://doi.org/10.1111/bij.12581

Peters, D.P.C., Loescher, H.W., SanClements, M.D. & Havstad, K.M. 2014. Taking the pulse of a continent: expanding site-based research infrastructure for regional- to continental-scale ecology. *Ecosphere* 5, art29. https://doi.org/10.1890/ES13-00295.1

Pocock, M., August, T., Rolph, S., Mondain-Monval, T., Jarvis, S., Burkmar, R., Baird, K., Pateman, R., Dyke, A., Wright, E., McInerny, G. & Turkay, C. (n.d.) The DECIDE Recorder Tool: Recording nature where it matters [Web application].

Pocock, M.J.O., Logie, M.W., Isaac, N.J.B., Outhwaite, C.L. & August, T. 2019. Rapid assessment of the suitability of multi-species citizen science datasets for occupancy trend analysis. https://doi.org/10.1101/813626

Pockock, M.J.O., Harvey, M.C. & Harrower, C.A. 2023. The 'Targeting Revisits Map' and evaluation of its impact on recorder behaviour. JNCC Report 729. JNCC, Peterborough, ISSN 0963-8091. https://hub.jncc.gov.uk/assets/19f32565-f628-409e-b732-e72a9ae6df28

Reichert, B.E., Bayless, M., Cheng, T.L., Coleman, J.T.H., Francis, C.M., Frick, W.F., Gotthold, B.S., Irvine, K.M., Lausen, C., Li, H., Loeb, S.C., Reichard, J.D., Rodhouse, T.J., Segers, J.L., Siemers, J.L., Thogmartin, W.E. & Weller, T.J. 2021. NABat: A top-down, bottom-up solution to collaborative continental-scale monitoring. *Ambio* 50, 901–913. https://doi.org/10.1007/s13280-020-01411-y

Ripple, W.J., Wolf, C., Newsome, T.M., Galetti, M., Alamgir, M., Crist, E., Mahmoud, M.I. & Laurance, W.F. & 15,364 scientist signatories from 184 countries. 2017. World Scientists' Warning to Humanity: A Second Notice. *BioScience* 67, 1026–1028. https://doi.org/10.1093/biosci/bix125

Shepherd, K.D., Shepherd, G. & Walsh, M.G. 2015. Land health surveillance and response: A framework for evidence-informed land management. *Agricultural Systems* 132, 93–106. https://doi.org/10.1016/j.agsy.2014.09.002

Sparrow, B.D., Edwards, W., Munroe, S.E.M., Wardle, G.M., Guerin, G.R., Bastin, J.-F., Morris, B., Christensen, R., Phinn, S. & Lowe, A.J. 2020. Effective ecosystem monitoring requires a multi-scaled approach. *Biological Reviews* 95, 1706–1719. https://doi.org/10.1111/brv.12636

Staley, J.T., Redhead, J.W., O'Connor, R.S., Jarvis, S.G., Siriwardena, G.M., Henderson, I.G., Botham, M.S., Carvell, C., Smart, S.M., Phillips, S., Jones, N., McCracken, M.E., Christelow, J., Howell, K. & Pywell, R.F. 2021. Designing a survey to monitor multi-scale impacts of agri-environment schemes on mobile taxa. *Journal of Environmental Management* 290, 112589. https://doi.org/10.1016/j.jenvman.2021.112589

Stauffer, N.G., Duniway, M.C., Karl, J.W. & Nauman, T.W. 2022. Sampling design workflows and tools to support adaptive monitoring and management. *Rangelands*, 44, 8–16. https://doi.org/10.1016/j.rala.2021.08.005

Sugai, L.S.M., Silva, T.S.F., Ribeiro, J.W. & Llusia, D. 2019. Terrestrial Passive Acoustic Monitoring: Review and Perspectives. *BioScience* 69, 15–25.

Thomaes, A., Barbalat, S., Bardiani, M., Bower, L., Campanaro, A., Fanega Sleziak, N., Gonçalo Soutinho, J., Govaert, S., Harvey, D., Hawes, C., Kadej, M., Méndez, M., Meriguet, B., Rink, M., Rossi De Gasperis, S., Ruyts, S., Jelaska, L.Š., Smit, J., Smolis, A., Snegin, E., Tagliani, A. & Vrezec, A. 2021. The European Stag Beetle (Lucanus cervus) Monitoring Network: International Citizen Science Cooperation Reveals Regional Differences in Phenology and Temperature Response. *Insects* 12, 813. https://doi.org/10.3390/insects12090813

Thomsen, P.F. & Willerslev, E. 2015. Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation*, Special Issue: Environmental DNA: A powerful new tool for biological conservation 183, 4–18. https://doi.org/10.1016/j.biocon.2014.11.019

Tulloch, A.I.T., Possingham, H.P., Joseph, L.N., Szabo, J. & Martin, T.G. 2013. Realising the full potential of citizen science monitoring programs. *Biological Conservation* 165, 128–138. https://doi.org/10.1016/j.biocon.2013.05.025

Turrini, T., Dörler, D., Richter, A., Heigl, F. & Bonn, A. 2018. The threefold potential of environmental citizen science - Generating knowledge, creating learning opportunities and enabling civic participation. *Biological Conservation* 225, 176–186. https://doi.org/10.1016/j.biocon.2018.03.024

Uchida, K., Blakey, R.V., Burger, J.R., Cooper, D.S., Niesner, C.A. & Blumstein, D.T. 2021. Urban Biodiversity and the Importance of Scale. *Trends in Ecology & Evolution* 36, 123–131. https://doi.org/10.1016/j.tree.2020.10.011

Wright, A.D., Campbell Grant, E.H. & Zipkin, E.F. 2022. A comparison of monitoring designs to assess wildlife community parameters across spatial scales. *Ecological Applications* 32, e2621. https://doi.org/10.1002/eap.2621

WWF, 2022. Living Planet Report.

Young, N.E., Stohlgren, T.J., Evangelista, P.H., Kumar, S., Graham, J. & Newman, G. 2012. Regional data refine local predictions: modeling the distribution of plant species abundance on a portion of the central plains. *Environ Monit Assess* 184, 5439–5451. https://doi.org/10.1007/s10661-011-2351-9

Appendix 1: 'Tracking the Impact' Case Study

Summary

'Tracking the Impact' is a landscape-scale conservation initiative in southern England which includes landscape-scale monitoring of birds, butterflies, and plants, delivered through citizen science, making use of the protocols from national monitoring schemes to monitor the state of biodiversity.

Benefits and successes of the project after the first field season include:

- Engaging 125 volunteers.
- Running training programmes to upskill volunteers.
- Provision of landscape-scale data.
- Provision of cross-taxa data.

Key lessons learnt from the project include:

- More work is needed to understand how to increase the diversity of recorders.
- It is important to ensure that projects such as this can leave a legacy and can be sustained long-term.
- Training opportunities should be appropriately balanced between online and in-person.
- Accessibility can be a barrier to participation, whether from a physical perspective (e.g. needing to access private land) or a financial perspective (e.g. volunteers not being able to afford to reach survey destinations).
- It is important to ensure alignment with other monitoring schemes and to not 'reinvent the wheel'.
- Monitoring should be designed to meet the needs of relevant use cases.

It was considered that it would be feasible to replicate the project elsewhere and scale it up, if:

- Resources were available and optimised.
- Local differences, such as how to define project boundaries, could be agreed on.
- Relationships with volunteers and partner organisations could be maintained effectively at the scale proposed.

Background and context

This case study is based on an example of a successful multi-taxa citizen science project that has taken place at a landscape scale. It summarises the project and what it measures, explores key benefits the project brings and challenges it has faced, and assesses the potential for the format to be replicated in other areas and at other scales. The project was recently awarded the 2022 BTO Marsh Award for Local Ornithology recognising its potential as replicable landscape-scale monitoring model. It is therefore hoped that this information could help inform the development of similar landscape-scale monitoring projects elsewhere. For example, it may be of interest to work on establishing Natural Capital and Ecosystems Assessment (NCEA) citizen science protocols, in particular in relation to the exploration of Regional Hubs.

What is the 'Tracking the Impact' project?

'Tracking the Impact' is a landscape-scale conservation initiative covering about a third of the Chilterns Area of Outstanding Natural Beauty (AONB). It is a five-year, National Lottery Heritage Fund project, run by the Chilterns Conservation Board in partnership with Butterfly Conservation, the British Trust for Ornithology (BTO), the UK Centre for Ecology and Hydrology (UKCEH), Buckinghamshire and Milton Keynes Environmental Record Centre, Berkshire (BMERC), Buckinghamshire and Oxfordshire Wildlife Trust (BOWT) and Plantlife.

It consists of two parts:

A. On-the-ground conservation, delivered through farmer clusters. This includes creating, managing, and connecting habitat across areas belonging to different landowners within the landscape.
B. Landscape-scale monitoring, delivered through citizen science. This aims to both help show whether the on-the-ground conservation activities have made a difference, and to upskill a new generation of citizen scientists with species identification and survey skills.

This case study primarily focuses on the latter part of the 'Tracking the Impact' project, to understand whether similar landscape-scale citizen science projects could be rolled out elsewhere.

What does it measure?

'Tracking the Impact' records bird, butterfly, and plant species across 50 x 1 km squares within a defined project area in the Buckinghamshire Chilterns.

Many conservation projects monitor their actions at a site level. However, this makes it difficult to understand holistic impacts of the conservation across the wider landscape. Data are also available at a national scale based on surveys such as the Breeding Bird Survey (BBS), the UK Butterfly Monitoring Scheme (UKBMS) and the National Plant Monitoring Scheme (NPMS), which are undertaken by citizen scientists. However, these schemes are designed to work at a national scale, so do not have data from enough sample points within an area of the size that the 'Tracking the Impact' project is interested in to be representative. There is therefore a data gap for understanding trends at a landscape scale, which the project is aiming to address.

The approach taken by 'Tracking the Impact' makes use of the protocols from the national BBS, UKBMS and NPMS schemes. This ensures that the data collected are based on a robust method that is known to be tried and tested, and that is recognised as an industry standard. It also means it is replicable, so the same process could be undertaken anywhere else, and all the data could feed into a centralised system. Building on current schemes also means there is access to back-end support systems that are already built, and centralised training from the national schemes' websites, maximising the efficiency of resource required. These three taxa were selected as they were seen to be the most accessible for citizen scientists to record, were typical of the dominant habitat types in the landscape and were a robust proxy for wider trends.

BMERC divided the project area into 1 km squares. Any square falling in what they considered to be an urban area or with more than a third of its area falling outside the project area was excluded. Fifty squares of those remaining were selected as sample squares (see Figure 2). This was achieved using a semi-random generator, which allowed for inclusion of a representative balance of woodland compared to farmland. Fifty was selected by the project's steering group as the number of squares to use to ensure enough of the total project area was covered (20%) to be statistically valid and enable robust inferences of trends to be made across the full project area from the sample data.

With fifty squares and three schemes, there are 150 volunteering opportunities in total. New volunteers are sent all unclaimed squares and can select the square that would be easiest for them to take responsibility for monitoring. Many volunteers only participate in one of the three schemes, but there is some crossover.

This monitoring will allow for estimates of trends in bird, butterfly and plant populations and diversity over time. The steering group hope to use this as a proxy for understanding landscape-scale change.

Less structured monitoring is also undertaken, including:

- The Farm Walkover Bird Survey project (4 years across 18 farms with a team of 14 volunteers, data uploaded directly onto BirdTrack)
- Amphibian/Reptile surveys (2 years across 18 farms with a team of 20 volunteers, data uploaded into National Reptile Survey)
- A bird ringing project (3 years across 5 farms with a team of 7 volunteers, data entered into DemOn). This was also associated with a BTO approved Corn Bunting colour ringing project.
- A series of Rapid Habitat Assessments across 3 chalk grassland sites (with a team of 30+ volunteers)

It would be possible to expand the same structured approach to cover other taxa with similar national schemes if there was the relevant interest, expertise, and resource available.

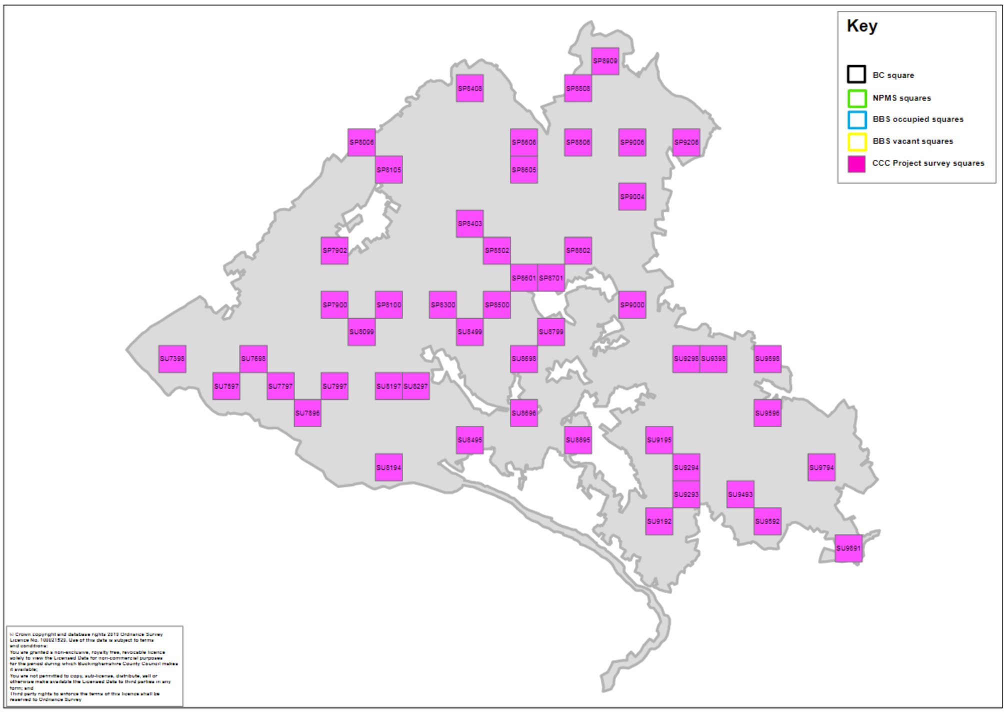

Figure 2. A map of the 'Tracking the Impact' project's area, showing the 50 squares semi-randomly selected as survey sites. Figure provided courtesy of Tracking the Impact project, reproduced with permission.

How do the data collected relate to national schemes?

The data collected by citizen scientists as part of the ‘Tracking the Impact’ study do not feed into national schemes directly. This is because the national schemes rely on a standardised and systematic process to select squares from which data should be collected. The inclusion of additional squares that were not randomly selected as part of the national protocol would bias national results, especially if all originated from within a small geographic area. The data do sit with the authorities running the national schemes (BTO, Plantlife and Butterfly Conservation) and add value to the dataset for potential use in more *ad hoc* analyses, but do not count towards the annual trends used in official national statistics. As the data are based on the same robust methods, comparative analyses could be undertaken to show the difference between landscape and national trends (for example through capacity being developed as part of the NERC Knowledge Exchange Fellowship: Bringing the data revolution to nature recovery).

Some of the squares used within ‘Tracking the Impact’ match up with squares used in national schemes. For example, 11 of the 50 ‘Tracking the Impact’ squares are also used within the national BBS scheme. This was done to ensure that the landscape scheme and national scheme could work symbiotically, with the drive to increase volunteer numbers in the local area contributing to greater coverage of nearby squares requiring surveillance for the national schemes, rather than drawing current national scheme volunteers away from the national schemes to participate in the local scheme instead.

The roll-out story

The project was lined up and ready to roll out for a first field season to take place in summer 2020. Covid restrictions greatly affected what was possible to achieve in the first year, but summers 2021 and 2022 were able to go ahead largely as planned. In summer 2021, data were collected in 55% of the 150 selected squares across the three surveys. As of summer 2022, 80% of squares had been allocated to a volunteer (including all 50 squares for birds, 38 squares for butterflies and 36 squares for plants). “Allocated to a volunteer” means that someone has agreed to do it, but the number of squares that will be completed may be lower than this. The project aims for 90% coverage by the final year.

Benefits and successes of the project

Volunteer engagement and upskilling

As well as being on track to achieve its aim of having 90% of squares covered, the project has so far engaged 125 volunteers. Of these, 70–80 are participating in the surveys, while the others have only been involved in the species identification training that the project has provided. Two species identification training groups have been run for each of the taxa (birds, butterflies, and plants). These were made up of four sessions in the field and four sessions run remotely. Training was aimed at those who wanted to develop their ID skills before thinking about picking up a survey square in future years.

Three training courses have also been delivered covering detailed survey methods across the three surveys. Training was aimed at volunteers who were confident in their species identification skills but had not carried out a formal survey before.

Another key benefit of the project is around volunteer motivation. Feedback from volunteers has shown that the high level of connection to local strategic work is a strong motivating factor for the volunteers who have got involved. There has been a lot of cross-over between volunteers who have got involved with the monitoring and those who have got involved in the on-the-ground conservation work. Indicating that those who get involved with monitoring learn more about the local area and want to help conserve it, and those volunteering for conservation activities recognise the important of monitoring to help tell the story of the local area and how well these conservation activities are working. Another key motivating factor is understanding the impact that an individual's volunteering has had and how their data are being used, especially in cases where it is used to inform planning decisions. Presenting data back to individual volunteers as feedback is not yet possible as part of the project but has been flagged as a potential motivating factor that could be implemented in future. Volunteers were not found to be strongly motivated by extrinsic factors such as certificates.

Value-add of providing landscape-scale data

At the time of writing (summer 2022), the project has only collected data from one full field season. It is therefore not yet possible to perform any analyses of landscape-scale trends or comparisons between the landscape and the national average. However, these are key benefits that the project expects to deliver in future years. It is expected that 5–10 years of data would be needed before true trends could be assessed, with confidence that differences are not due to factors like fluctuations between a wet summer and a dry summer.

A specification for the presentation of data, mapping and analysis is being worked up for delivery in winter 2022/23 to present results at a landscape, taxa, and individual square basis. It is hoped this will provide a consistent model for aggregated data to be reported on an annual basis.

The data available now are, however, useful for engaging landowners. For example, they are being used to demonstrate what is currently found within a particular farm, to encourage farmers' and landowners' enthusiasm about conservation activities on their land.

Ideally, the Chilterns Conservation Board hope that data from the project, and from similar local projects taking place elsewhere, could be considered in – and could add value to – reporting requirements for key policies that will affect the local area, such as part of the ELM (Environmental Land Management) funding settlements for farmers, for farmers to use in the Farming in Protected Landscapes Scheme, and to help with LNRS (Local Nature Recovery Strategy) reporting.

In time, they hope to be able to produce a report for the project area, analysing results for each of the three taxa being monitored. Without data from the project, this would not be possible. Only two or three national scheme squares for the butterfly and plant schemes and eleven squares for the BBS are found within the project area, which is not enough to make any kind of robust landscape-scale assessment. Raw data from the project are not currently publicly available to download, but they are available to the project coordinator who hopes to use them to produce and publish reports.

Provision of comparable cross-taxa data

As the 'Tracking the Impact' scheme uses the same fifty squares for each of the three surveys (unlike the national schemes), it will also be possible to investigate cross-taxa relationships. This may allow for more direct analyses of questions such as whether all three are moving (e.g. shifting population trends or range shifts due to climate change or other factors).

Challenges and lessons learnt

Improving volunteer engagement and upskilling

Whilst the numbers of volunteers the project has been able to engage with have been high, there remains a lack of diversity in the types of volunteers who are being engaged. Although there is a balanced gender ratio, two thirds of participants are white, middle class and retired. There is a lack of engagement with a more ethnically diverse audience or with younger people. This is likely a reflection of the wider problem of a lack of diversity across monitoring schemes more generally, having advertised through channels that would target this audience, such as conservation-focused social media pages, conservation NGOs, existing recording groups and presentations at local interest groups. It is also likely linked to wider issues around time and availability associated with these demographic groups (especially those of working age). This means that many existing surveyors are picking up more opportunities or developing skills in a different taxon, but there are fewer examples of truly new engagement of those entering the citizen science community for the first time. In general, local schemes are more likely to be able to get to local people on the ground and improve diversity than national schemes are, but there is a clear need for further work in this area to build understanding of how best to do so.

The lack of youth participation is an area that the Chilterns Conservation Board has been trying to address through other projects. They have worked closely with BTO youth representatives, who have explored some of the key barriers to joining a monitoring scheme that young people face and designed a new two-year project called New Shoots to help break these down. The project is aimed at 15–18-year-olds who already have an interest in nature and conservation but want to take the next step and learn more. This involves carrying out surveys of birds, butterflies, plants, and chalk stream invertebrates during summer, and assisting with restoration and scrub bashing work in winter. It also involves events such as behind the scenes tours of RSPB (Royal Society for the Protection of Birds) reserves and a bursary scheme that assists young people with purchasing binoculars, reimbursing travel, and paying for bird ringing courses. This will not produce new recorders straight away, but it will provide the foundations for a new recorder cohort for the future.

Ensuring a legacy

The project is half-way through its funded period, with two and a half years remaining. The project recognises that lottery funding is not sustainable longer term. There is therefore a need to make sure there is a project legacy, especially as it takes five or ten years to get meaningful data and identify trends. A key lesson learnt is therefore to integrate thinking from the beginning about how the project can leave a legacy and be sustained long-term.

Chilterns Conservation Board are planning to expand the project into a further 25 x 1 km squares as part of a wider Thames Water funded Smarter Water Catchment project in the River Chess catchment – an adjoining landscape, offering a sense of scale and replicability.

Covid

The project took several key lessons away from experiences with Covid restrictions. When Covid started, training sessions and engagement switched to online, which was not an option they had considered making use of before. This included a whole summer of online identification and survey methods training sessions, and an engagement session with Chris Packham. Based on the popularity of this format, they have retained a hybrid system going forwards. Online training (and recordings) allows for much greater scalability and centralisation of the training sessions in cases where the same information is relevant across localities. For example, the online event with Chris Packham would have cost the same if it were watched by ten people or ten thousand people. Projects should therefore consider how best to offer a balanced and efficient suite of training options. However, local projects with in-person sessions do have advantages in terms of providing a greater ability to tailor teaching to the specific species assemblages associated with the local landscape and providing practical hands-on experience of carrying out surveys, so it is important that not all training is replaced with online training.

Accessibility

A key challenge faced particularly in terms of engaging a more diverse audience, is accessibility. For example, if someone living in a city without a car wants to participate, they are unable to access a survey that may be 25 miles or even less from their home. Biasing squares so some of them are located on mainline bus routes could be one way to help break down this barrier. It would need to be shown that this would not bias results. However, there would still be a cost associated with this (or with fuel prices for those who do have access to a car) which can form another barrier to access.

Physical access to private land can also be an issue in some cases, particularly with regards to plant surveys. The Chilterns Conservation Board have helped in some cases with identifying farmers for volunteers to ask for permission to survey on their land, but this is a time-consuming process.

Reducing recorder bias

The project asks volunteers to self-assess their species identification skills, through questions such as “How many bird species could you confidently identify?” However, it remains a challenge to understand the extent to which skill level differs across squares. Nonetheless, this is the same situation as with the national scheme data, which are considered robust and well-trusted. Any volunteers that do not have a skill level appropriate to complete their own square are offered alternative ways to get involved, such as through the training sessions. Consistency of effort within the same square is considered more significant than potential differences between squares.

Another possible source of bias is which squares were selected by volunteers. Squares were not prioritised in terms of which were more important to be completed in any way, and so volunteers may have ignored squares with no public footpaths or preferentially selected squares closer to population centres. Once again, this is a well-established caveat of relevance to the national schemes as well, which may even be less significant in the example of the ‘Tracking the Impact’ project due to the very high coverage of squares overall.

Avoid reinventing the wheel

Another key lesson learnt from the project is the need to avoid dividing recording effort. Local projects will not add value if they dilute or confuse schemes that are already taking place. It is necessary to ensure that local data feed into national repositories and take complimentary approaches that allow for robust comparisons.

Ensure monitoring is designed to be tied in with relevant use cases

Another key challenge has been trying to ensure that the data produced will be relevant to feed into specific use cases. Driving appetite for the need for monitoring before undertaking surveys will help to inform the design of the process, and to ensure that the data produced can be as useful as possible. This includes considerations around how the data can be used more widely, highlighting once again the need for alignment between local projects to enable comparisons. A single project with one prescriptive process will not be able to fulfil all use cases alone; it is likely that some flexibility would be needed, which could for example be provided through an adaptable package of tools. For example, from a farmer's perspective, having surveys done on 1 km squares of their farm is interesting, but it does not inform them about the rest of the farm, which would be necessary if wanting to use the data for purposes such as ELM assessments. This would require volunteers to take a more *ad hoc* approach to monitoring, which could be elicited from a different tool in the package.

Implications of wider thinking

Ensuring that the landscape-scale data collected is of use to inform wider thinking would be helpful. For example, 'Tracking the Impact' provides good data for the third of the Chilterns that it covers, but there are not enough data in the other parts of the Chilterns to inform wider thinking about the AONB overall.

Determining relative costs and benefits

Determining the relative costs and benefits of setting up a local project would also be useful to incorporate into planning and reporting. For example, BTO have recently analysed the carbon footprint of volunteers undertaking monitoring by comparing home address postcode with survey square locations, accounting for the reported transport methods between the two. This kind of analysis has not been undertaken for 'Tracking the Impact' but could be of interest to understand whether local schemes could offer a lower carbon footprint per unit of data collected due to potentially shorter travel distances.

Feasibility of replicating the project elsewhere

Through conversations with other chalk based AONBs around the potential for replicating the project elsewhere, the 'Tracking the Impact' project has found that there is definite appetite for landscape monitoring, but other bodies face a number of barriers to being able to implement it.

Resourcing requirements

One key barrier is the resourcing requirements that it takes to run such a project. Many organisations simply do not have the resourcing and staff to be able to run such a project, as monitoring is generally under-resourced.

The 'Tracking the Impact' project is run by a coordinator who spends an average of two days a week on the project over the year. This time is skewed to highest commitment in spring and lowest commitment in autumn. This current model may not be cost optimal (i.e. it could be possible to coordinate squares with less time commitment). However, feedback from volunteers has been that they strongly value having someone who puts the time in to answer queries and generate enthusiasm. If things became too automated and centrally generated, volunteers would lose some of the excitement they feel around the local nature of the project.

In addition to the time spent by the project coordinator, there is also a small budget (approximately £4,500 per year) for delivery of training and one-off fees for adaptations to national online databases. It may also be possible to optimise this if scaling up, for example by including more people on training courses whilst maintaining the same costs or centralising training courses through online delivery, but again smaller courses give a more personal and practical experience of surveys. This scaled-up model is being tested in the roll out of the project into the River Chess catchment.

When the project was being set up, a few days of work each also went in from the local record centre, who generated the sample squares, and BTO, BC and Plantlife, who put in time to set up the database being used to collect the data.

Geographic differences

No unique factors about the Chilterns that make it an easier or more difficult area to run this kind of project compared to anywhere else in the country were identified. The only difference is having the resource and the will to make it happen. However, if implementing similar projects elsewhere, different areas might have specific interests in different taxa. One approach to account for this could be to have a common set of generic taxa that are monitored in any local project (e.g. birds, butterflies, and plants), with a flexible ‘menu’ of additional taxa that could be added onto local projects where they are of particular interest (e.g. the *ad hoc* reptile recording being undertaken as part of the ‘Tracking the Impact’ project). Understanding how this would work on a geographic level, as well as on a policy level, would be a challenge to consider if implementing similar projects elsewhere.

Geographic boundaries

Another key challenge faced raised in conversations about replicating the project in other chalk based AONBs is around how to define project areas. There is no real agreement around whether it is most useful and practical to tie in project boundaries with ecologically relevant areas (such as river catchments), administratively relevant areas (such as counties or protected landscapes), areas related to potential funding streams (such as those covered by Wildlife Trusts) or those defined by relevant policies (e.g. LNRS areas).

Feasibility of scaling up the project area

If attempting to run a similar project over a larger area, a few key considerations would need to be taken. There will likely be a trade-off between the scale at which resources and costs are optimised and the scale at which relationships with volunteers and partner organisations are optimised.

Resource optimisation

It is thought that the project coordination resource would scale well (i.e. taking on the coordination of double the number of squares would not require double the amount of time). Taking on an area the size of the whole of the Chilterns would be estimated to be a full-time role.

Relationship with volunteers

Feedback from volunteers suggests that the small scale of the project is a key motivating factor. They feel that the local area is ‘their patch’ and they therefore wish to do their part to monitor, protect and care for it. There is a risk that scaling the project up to an area that dissociates volunteers from their sense of local action could lead to reduced participation.

Relationship with partner organisations

Scaling the project up would also require engagement and coordination with a larger number of project partners, which may lead to additional resource requirements. For example, if scaling up 'Tracking the Impact' to cover the whole of the Chilterns, this would involve working across four Local Environmental Record Centres and four county boundaries, instead of just the one. Similarly, being small in scale has allowed the project team to get to know local landowners and make use of direct engagement.

Could this model be used to inform development of other landscape-scale monitoring programmes, such as the NCEA citizen science protocols?

This case study highlights several factors that would need to be considered when developing landscape-scale monitoring and local citizen science protocols. These factors include the need for new monitoring activities to ensure they build on and align with existing schemes, the importance of local engagement and relationship building, and an indication of the scale of likely resource required. Whilst too early in the project cycle to draw firm conclusions, it appears that the model used in 'Tracking the Impact', based on local coordination of more data-intense versions of recognised national monitoring schemes, will prove to be a highly effective way of engaging volunteers and gaining data at the scale it has been implemented at in this project. The approach is likely to be scalable given sufficient resource, but if trying to implement across very large areas may have limitations in terms of losing local ownership.