JNCC Report

No. 440

Designing a new plant surveillance scheme for the UK

Kevin Walker, Trevor Dines, Nicola Hutchinson and Stephen Freeman

September 2010

For further information please contact:
Joint Nature Conservation Committee
Monkstone House, City Road
Peterborough, PE1 1JY, UK
www.jncc.gov.uk

Preface

This report summarises the work undertaken to design an 'ideal' UK vascular plant surveillance scheme that complements other species and habitat surveillance in providing an understanding of the overall state of the UK environment. The need for such a scheme was identified in the UK Terrestrial Biodiversity Surveillance Strategy (UKTBSS) which seeks to ensure that surveillance is undertaken strategically; integrating needs for evidence and making best use of existing surveillance data (JNCC 2009). The research in this report combines to outline the parameters and deliverables of a new plant surveillance scheme.

The work was undertaken by a partnership of organisations with unparalleled experience of the design and delivery of surveillance and recording schemes in Britain. Drawing on expertise from within this partnership, the scheme was designed to be simple, low cost and achievable by volunteers, to use only a subset of (target) species, to include simple structural (vegetation/habitat) measures that could be collected at low cost, and to allow repeatable sampling within a maximum of a half a day of field work (given no access difficulties).

Seven methods of sampling vegetation were examined in detail, all based on surveying 1km squares. From these, three design options are recommended for subsequent field-testing. Recommendations are made regarding the selection of target species, the classification of habitats, the geographic stratification of the squares, the collection of a key set of vegetation and habitat parameters, the optimal time to survey, and the statistical implications that need to be considered for the recommended methods. Additionally, the relative costs, opportunities and constraints of implementing surveillance scheme programmes were considered by the project team, as well as the implications of delivering such a scheme with a network of volunteers.

Acknowledgments

The authors would like to thank members of the discussion group for their valuable comments and advice throughout the course of this project, in particular Michael Braithwaite (BSBI), Mick Crawley (Imperial College), Jenny Duckworth, Bob Ellis (BSBI), Alison Johnson (BTO), Quentin Groom (BSBI), Alex Lockton (BSBI), David Noble (BTO), Tim Pankhurst (Plantlife), David Pearman (BSBI), Ellen Pisolkar (Plantlife), Chris Preston (CEH), David Roy (CEH) and Simon Smart (CEH).

We would also like to thank Chris Cheffings, JNCC, for her support and guidance, NERC Centre for Ecology and Hydrology for access to the Environmental Change Network and Countryside Survey datasets which were used to derive some of the figures in Chapter 2 and Gwynn Ellis (BSBI) for data on BSBI membership used to produce Figure 4.4.

Contents

1 Introduction and background

The main aim of this project was to design an 'ideal' UK vascular plant surveillance scheme that complements other species and habitat surveillance in providing an understanding of the overall state of the UK environment.

This aim was set in the context of Objective 1 of the UK Terrestrial Biodiversity Surveillance Strategy (UKTBSS) which seeks to ensure that surveillance is undertaken in a strategic manner, integrating needs for evidence and making best use of existing surveillance data (JNCC 2009). Objective 1 of the UKTBSS requires the measurement of status and trends of a framework of habitats, species, and their ecosystem functions, sufficient to inform the delivery of the outcomes required by UK and country biodiversity strategies.

Plants are the fundamental building blocks of all habitats and ecosystems, providing food and shelter for other wildlife, so it is essential that we assess the health of the natural environment using information on plant diversity, not least to provide indications of emerging environmental problems as well as positive signs of sustainable recovery. Establishing a robust plant surveillance scheme will vastly improve the UK's ability to report on and respond to the state of the natural environment.

1.1 Biodiversity policy and strategy

Biodiversity policy is a devolved matter in the UK with country biodiversity strategies providing the primary delivery frameworks. Regardless of devolution the range of delivery mechanisms available within each country can broadly be grouped into species recovery programmes, protected site networks, agri-environment and forestry schemes and wider countryside measures such as habitat restoration and landscape permeability. Whilst surveillance should be an integral part of these delivery mechanisms, this is not consistently applied across each of the UK countries, and certainly not with respect to plant populations.

There is a need to understand the influence of existing policies and inform the changes required to enhance the health of the natural environment. Issues that require particular detection include the impacts and intensity of eutrophication (agricultural run-off, transport and industry), habitat fragmentation, habitat management, changes in agricultural crops, factors related to climate change and non-native invasive species. Similarly there is a need to understand the positive effects, or otherwise, of measures supporting population resilience and connectivity (such as landscape-scale conservation, positive land-use planning, etc), site protection (including under the NERC Section 40 biodiversity duty), and wider countryside initiatives (e.g. agri-environment schemes). A plant surveillance scheme can also aid the assessment and development of issues, policies and measures.

Table 1.1 outlines proven relationships between environmental drivers and plant traits. With the exception of mean July temperature of the realised species distribution, the known traits (corresponding explanatory variables) have established relationships with key drivers and provide robust indicators of the impacts of environmental pressures. The ecological basis of these responses are well understood, although for climate warming, the impacts of which are only just beginning to emerge, suggestions are predictive and require further research. Other drivers of change may be indicated through other Ellenberg values (such as light, moisture and acidity) but these relationships have not yet been proven through existing datasets.

Table 1.1. Key environmental drivers with proven relationships to plants as established through analyses of existing GB surveillance datasets (e.g. Smart *et al* 2005; Braithwaite *et al* 2005; Morecroft *et al* 2009).

Driver	**Ecological interpretation**	**Explanatory variable (established trait)**	**Relationship with driver**
1. Exposure to greater nutrient loads – atmospheric, terrestrial, aquatic (eutrophication)	Increased availability of soil macro-nutrients leading to increased above-ground competition	Ellenberg fertility value (N-value)	Increase and decrease
2. Reduced disturbance / under-grazing in specific habitats/areas (succession)	Reductions in the removal of standing biomass leading to increased above-ground competition	Canopy height Localised area of species occupancy reduced/lack of species regeneration/change in % cover of vegetation composition	Increase and decrease
3. Increased sheep and deer grazing / over-grazing (disturbance)	Increased removal of standing biomass combined with selective removal of palatable species (especially in the uplands)	Canopy height Species composition	Decrease and increase
4. Climate warming (temperature)	Increased warming	Mean July temp/Major Biome Vegetation composition and abundance of species	Increase and decrease

Beyond the key environmental drivers listed in Table 1.1 there are a wider range of issues, policies, strategies, indicators and mechanisms for which data from a plant surveillance scheme can provide an evidence base and trend information (see Tables 1.2 and 1.3). In England, the emergence of a Natural Environment White Paper is a recent development which will equally require measurement of changes in the environment to which a plant surveillance scheme can contribute.

Table 1.2. Longer list of proven and unproven environmental issues and drivers linked to policy, legislation, strategy and incentive

Issues/drivers of change	**Policies/legislation/regulation/strategies/ mechanisms/incentives**
Exposure to greater nutrient loads: run-off, transport, industry, direct agro-chemical application (eutrophication)	Water Framework Directive (River Basin Management Plans) Environmental Liability Directive Water Resources Act CAP/Agri-environment schemes National Emissions Ceilings Directive Air Quality Strategy Integrated Pollution Prevention and Control Directive
Reduced disturbance/under-grazing (succession)	CAP/Agri-environment schemes BAP (habitat management)
Increased (over) sheep and deer grazing (disturbance)	Deer Initiative CAP/Agri-environment schemes BAP (habitat management)
Changes in agricultural and forestry land management/drivers (leading to changes in management, habitat fragmentation, degradation and loss)	Forestry strategies: e.g. England's Trees Woods and Forests (ETWF) Delivery Plan 2008-2012; restoration of open habitats policy Water management strategies Environmental Liability Directive CAP/Agri-environment schemes Woodland grant schemes (England)
Climate warming (temperature) – and also increase in extreme events – droughting, heavy precipitation , sea level rise, flooding (habitat/species loss/change, changes to phenology)	Climate Change Act Biodiversity adaptation policy
Changes in infrastructure: housing, transport, coastal defence, tourism/recreation, hydroelectric (leading to habitat fragmentation, habitat degradation, habitat loss)	Planning Policy Statements Section 40 NERC Act; National Indicator 197 Green Infrastructure strategies EU Habitats Directive
Coastal erosion (habitat loss)	
Extraction: minerals (habitat loss)	Section 40 NERC Act ROMPS; other minerals policies
Other human impacts: trampling, interference, arson (habitat degradation)	WCA Sch8; EU Habitats Directive
Non-native invasive species (competition, habitat degradation) Plant pathogens GM Crops	WCA Sch9; ban on sale
Conservation land management & landscape scale conservation/connectivity	Global Strategy for Plant Conservation / Plant Diversity Challenge BAP Country Biodiversity Strategies & indicators (see below) Protected site networks
Exploitation of species	WCA Sch 8 Sustainable use policies

Table 1.3. Country biodiversity strategies and indicators which data and trends from a UK-wide plant surveillance scheme could contribute to

Strategy	Indicator (as of March 2010)
UK Biodiversity Indicators in Your Pocket	2. Plant diversity 10. Ecological impacts of air pollution 11. Impact of invasive species 3. Status of UK Biodiversity Action Plan Priority Species
England Biodiversity Strategy	A4 Trends in plant diversity in fields and field margins in England F4 Trends in woodland plant diversity in England W4 Trends in plant diversity on river banks and stream sides in England A3 Status of farmland Biodiversity Action Plan priority species and habitats in England F3 Status of woodland Biodiversity Action Plan priority species and habitats in England H3 Status of Biodiversity Action Plan priority species and habitats in England W3 Status of water and wetland Biodiversity Action Plan priority species and habitats in England
Wales Environment Strategy	19: The loss of biodiversity has been halted and we can see a definite recovery in the number, range and genetic diversity of wildlife, including those species that need very specific conditions to survive. 19c: Indicators to illustrate range and genetic diversity (to be developed) 20: The wider environment is more favourable to biodiversity through appropriate management, reduced habitat fragmentation and increased extent and interconnectivity of habitats 20c: Additional indicators to be identified following completion of research into biodiversity indicators (to be developed)
Scottish Biodiversity Strategy	S01 Status of biodiversity action plan (BAP) priority species S06 Vascular plant diversity S07 Woodland diversity
Northern Ireland Biodiversity Strategy	BD1: sites of local nature conservation importance BD2: Priority Habitats BD3: Priority Species ASSIs

1.2 Biodiversity assessment and plant surveillance

Whilst changes in the status of the UK flora have been well documented in distribution atlases from national to local level, the absence of a regular national plant surveillance scheme has meant that trends, associated with specific drivers of change, have been more difficult to establish.

Currently, the UK assesses progress towards national and international biodiversity targets using a series of indicators including an indicator on plant diversity. The plant diversity indicator has, to date, been assessed using data from the Countryside Survey, which runs every 7-8 years, and was recently supplemented with data from the BSBI Local Change project which compared data from 1987-88 and 2003-04 (Braithwaite *et al* 2006). However, the plant diversity indicator under-performs in comparison with the indicators for birds, butterflies and bats, all of which can report changes on a more frequent basis.

The Surveillance Framework of the UK Surveillance strategy sought to answer a number of surveillance requirements and identify gaps in current coverage. Under objective 1 of the UK Surveillance Strategy the questions asked were as follows:

- Do we have the functional combinations of semi natural habitat we require in the landscape and how is habitat changing?
- Is the quality of semi natural habitat sufficient to maintain its function and species diversity, and how is this changing?
- Are species across ecosystem functions, and dependent on different scales of habitat (from micro habitat to migratory), being sustained within the landscape and how are their populations changing?

The first two questions posed here are linked in terms of understanding functionality within habitats and landscapes, with habitat quality and species diversity used as measurements within this. Vegetation sampling provides a habitat quality measure in terms of structure and species composition. A strategically designed plant surveillance scheme will complement existing vegetation assessments, such as those carried out within the Countryside Survey, filling gaps and adding value to current measures of habitat quality and condition. Plant surveillance inherently goes beyond simple species lists, taking account of ground and site conditions, allowing inferences on multiple environmental factors such as the impacts of land management on within-habitat heterogeneity.

The third question from the Surveillance Framework under objective 1 seeks information on regular and reliable trends for species. To date in the UK, vascular plant surveillance that includes repeat measures has been carried out across a range of schemes which vary in temporal frequency and spatial design, including sample selection, size, type and survey method (see Table 1.4). Despite unequivocal evidence of the impacts of environmental pressures, such as eutrophication, from single plant-based studies carried out at a range of scales (e.g. Preston *et al* 2002; Smart *et al* 2005; Braithwaite *et al* 2006; Walker *et al* 2009; Keith *et al* 2009), it has been difficult to detect consistent trends from across the range of plant surveillance schemes in operation due to their variations. A more strategic approach to plant surveillance is therefore required.

Table 1.4. National surveillance schemes covering vascular plants

Common Plants Survey (CPS)	Led by Plantlife since 2000, recording 65 common plants annually in over 3000 small plots (within c. 550 1km random squares) to date. The survey method was designed to complement the Countryside Survey.
Countryside Survey (CS)	Led by CEH, running since 1978 at approximately 7year intervals, focused on a sample of c.600 1km stratified random squares representing all major habitat types, with plots of 4-25m^2.
Local Change (LC)	Led by BSBI and initiated in 1987. Surveyors record all species occurring within systematic grid of c. 880 tetrads (2km^2) every 15 years.
Environmental Change Network (ECN)	Multi-agency programme focused on 0.01-4 m^2 plots within12 terrestrial sites (began 1993) and 45 freshwater sites (began 1994). Repeated every 3-9 years, measurements are a range of physical, chemical and biological variables.

With a strategically designed annual plant surveillance scheme focused on common species in place, value would be considerably added to the UK's current assessment of the natural environment. For instance, the measures on habitat quality obtained would complement the monitoring routinely carried out on protected sites (e.g. Common Standards Monitoring) or through the BAP process (i.e. priority species and habitats), thereby providing a 'wider countryside' comparison. An annual plant surveillance scheme of the scope outlined in this report would similarly broaden and add weight to the pictures revealed through other long-term surveillance schemes such as the Breeding Bird Survey or Butterfly Monitoring Scheme, by being implemented at a comparable spatial and temporal scale.

A greater understanding of the health of plant populations within the UK is paramount to key policy and practice decisions. One example is the work of BICCO-NET which is seeking indicators of the impact of climate change on the natural environment and requires annual abundance data to assess changes relevant for climate change adaptation management. The ongoing National Ecosystem Assessment's (NEA) preliminary report states that "micro-organisms, fungi and plants play a role in underpinning all provisioning and regulating services whereas while vertebrate groups play a role in all cultural services, they only play an important role in 25% (2/8) of the provisioning and regulating services". As the emphasis of how and why we conserve the natural environment shifts more towards the delivery of ecosystem services this simple but powerful statement from the NEA illustrates how vital plants are in terms of underpinning the services provided by the natural environment, both in terms of the mechanics of ecosystem service delivery and direct ecosystem service provision.

The majority of the UK's terrestrial natural environment is subject to the influence of farming or forestry. Whilst there are numerous schemes focused on improving conditions for species and habitats currently the monitoring of the impact of these is sporadic, at an inadequate intensity and often taxonomically narrow in its focus. A plant surveillance scheme covering common species allows, as mentioned previously, for a health check at the wider countryside level and as such would help to fill gaps in understanding of the changes in extensive arable or forested landscapes.

The plant surveillance scheme described in this report is an excellent example of the Coalition Government's Big Society agenda, allowing communities to get involved with taking responsibility for monitoring the natural environment. Plant diversity is fundamental to all ecosystems and a civil society scheme to monitor trends in plant diversity will draw on and build skills such that individuals better understand and become concerned with highlighting areas of progress and concerns, leading to greater involvement in discussing solutions to enhance the natural environment and its services.

2 Considerations for surveillance scheme design

2.1 What measures

To meet the objectives of the UK TBSS objective 1 and track the effects of environmental drivers listed in Tables 1.1 and 1.2 a new surveillance scheme would need to provide robust measures for the following attributes:

- Habitat change;
- Habitat quality (including changes in species composition and diversity);
- Individual species trends (including nationally threatened or BAP species where there is sufficient replication across the surveillance network);
- Combined species trends in relation to functional/ecological groupings. These should provide clear 'signals' in relation to anticipated/unanticipated drivers.

Given that a new scheme is intended to be low cost (see below), a key aim would be to record a few simple measures that can be used to report against all these attributes more frequently and over a much greater geographic area than existing schemes. Although some of these variables are currently well reported (e.g. changes in habitat extent by Countryside Survey) a new plant scheme could provide complementary information that would either augment or extend the sampling domain for poorly sampled areas, habitats and species, in particular the quality of habitats in the wider countryside. More frequent measures of occupancy and abundance would also provide more frequent assessments of species trend, including for 'priority' species and other 'indicators' of habitat quality, in response to changing drivers. As part of the new scheme such assessments could include positive effects for example following the introduction of policy initiatives to enhance or restore habitats (e.g. through agri-environment schemes) as well as the negative effects of anticipated and unanticipated environmental drivers. In recording a core set of target species, representative of a broad range of plant traits, such an approach would also enable analyses of functional/ecological groupings to address the direction, nature and scale of these positive and negative environmental drivers.

2.2 Other considerations

JNCC recommended that an 'ideal' surveillance scheme should be designed to take into account the following points:

- the method should be simple, low cost and achievable by volunteers or professionals;
- the scheme should use only a subset of (target) species;
- the scheme should include simple structural (vegetation/habitat) measures that could be collected at low cost;
- completion of each 'sample' to be achievable in a repeatable manner within a maximum of half a day of field work (given no access difficulties).

It was recommended that five broad questions be considered:

1. What is the optimal set of species to target in order to provide good representation of a broad range of plant traits?
2. What are the benefits of a range of plot types and selection methods?

3. What is the optimal number of visits, bearing in mind that different species are more visible at different times of the year?
4. What is the optimal set of parameters to be recorded?
5. How many plots / transects need to be included in order to record significant change over 5-10 years?

Sections 3 to 8 of this report seek to address these questions.

Though the stated aim of the project was to design an 'ideal' plant surveillance scheme, it was also clear from these requirements and questions that the project should seek to enhance, rather than simply replace, existing plant surveillance programmes. Increasing complementarity between these, and possibly linking to surveillance schemes for other taxonomic groups, would be highly desirable, especially for potential funding bodies. This could be achieved though modifications and/or additional modules that increase the consistency of approach across existing schemes and maximize the explanatory power of the combined sampling framework. Such an approach has substantial benefits including increased sample replication, better use of existing surveillance data and well-established volunteer networks, and potential integration of results with other taxonomic groups.

Although there may be scope to employ 'low-cost' professionals to implement some aspects of a new plant surveillance scheme it is assumed that the majority of the work would be carried out by unpaid volunteers (such as members of the BSBI, Plantlife, The Wildflower Society and other wildlife organisations, as well as newly interested and motivated members of the public). To be sustainable the new scheme therefore needs to appeal to a wide range of skill levels, allowing beginners as well as experts, to contribute useful information that can be analysed and ultimately used to meet the objectives of the scheme. Wide participation would have the additional benefits of improving geographic coverage and sample sizes thereby making the findings more robust and increasing the numbers of habitats and species that could potentially be reported on. The need to ensure wide participation has therefore had a significant bearing on the overall design of the new scheme presented in this report.

Additionally it was felt by the project team that the relative costs, opportunities and constraints of implementing a surveillance scheme programme needed to be considered. Although not a requirement of the project, the practical delivery of different surveillance design options have been considered. Biodiversity surveys carried out by volunteers have a number of strengths including recorder enthusiasm and ownership, geographic scope, potential for multi-taxa recording, opportunities for skills enhancement (the latter recognized as an important requirement in the UK Terrestrial Biodiversity Surveillance Strategy Vegetation Sampling Workshop 2008) and the potential for surveys over an extended period of time. However, there are also considerable constraints, for example the need for organisational support, variable skill levels, the need for the survey to be engaging and enjoyable for volunteers, reluctance to record in particular ways, retention of interest and access issues. The latter is especially important for plants; since plants are relatively immobile, recorders have to visit the sites where they grow and gaining access permission places a very considerable barrier to recording for many volunteers.

2.3 Project methods

A project team was established to submit the original tender and has continued to work together throughout the duration of the project. An e-discussion group engaged additional participants from the BSBI and Plantlife in January 2010 and discussed a range of issues outlined in Annex 1. A workshop was held in February 2010 at which the project partners discussed a series of pre-prepared papers on the selection of species, possible field methods and parameters to record (see Annex 3 for the agenda and meeting note). The project partners provided advice and support to Plantlife and the BSBI throughout the project, with Stephen Freeman of CEH undertaking the statistical analysis presented in section 8.

The original questions posed in the original project specification form the structure of this report as follows:

1. What is the optimal set of species to target in order to provide good representation of a broad range of plant traits? *Section 3.*
2. What are the benefits of a range of plot types and selection methods? *Sections 4 and 5.*
3. What is the optimal number of visits, bearing in mind that different species are more visible at different times of the year? *Section 6.*
4. What is the optimal set of parameters to be recorded? *Section 7.*
5. How many plots / transects need to be included in order to record significant change over 5-10 years? *Section 8.*

3 Selecting target species

To ensure wide participation at all levels of expertise, it is important that a surveillance scheme allows volunteers to record a subset of target species. In addition, the need for an adequate number of primary samples (1km squares/monads) to be collected and for a rapid survey method (one day or less), suggests a list of target species would be highly desirable. This approach, however, needs to be balanced with the possibility for more experienced recorders to record all the species and taxa they encounter with the exception of genera such as *Euphrasia*, *Hieracium*, *Taraxacum*, *Rubus* and a few species and infraspecific taxa that are not recorded consistently even by experts. The need to have an option to record all species is important because it reduces the bias towards certain habitats and drivers that a predetermined set inevitably brings. By doing this it allows for the detection of unanticipated effects on species that are currently resilient to climate change, eutrophication, etc. On a more practical level it will also make the scheme more attractive to the most experienced surveyors who routinely record everything and would find a more restricted pool probably more time-consuming to record especially if it included 200-500 species.

The advantages of defining a subset of target species include:

- Makes the survey simpler and more enjoyable thereby appealing to a wider range of participants;
- It is less time consuming to record;
- Allows annual collection of data (even the most dedicated of volunteers would probably be unwilling to record all taxa every year);
- It significantly increases participation and therefore amount of data collected;
- It can reduce recorder error if focused on easy-to-identify species;
- It can be more sensitive to *known* drivers of change if links to environmental changes are well established/proven;
- Samples can be pre-selected on the basis of known distributions.

Disadvantages include:

- It assumes that there is an optimal set of species and that we could predict what these are in advance of surveillance;
- It could introduce bias because any selection would be influenced by our preconceptions of change;
- It compromises the completeness of surveillance and our ability to detect unanticipated changes in the future;
- It is inefficient from a recording and analytical perspective; better to record everything and analyse a subset of targets once changes have been established.

However, many of these disadvantages are overcome if the targets share the same traits (and variation in traits) as the whole sample, once an appropriate number of species are included in the sub-set. Therefore, the goal of species selection should be to be representative of the full range of traits, rather than extremes or key species (although these approaches are briefly considered).

3.1 Levels of recording based on ease of identification

Although a subset of target species can be employed, this does not mean that skilled surveyors would be excluded from recording all taxa. While an identical survey method would be employed by all participants, there could be several levels of recording depending on ability, botanical skills and the amount of time participants have available. A simple hierarchy would be:

Level 3	Record all species encountered (with the exception of some difficult genera)
Level 2	Record from a selection of c.300-400 easy- to moderately-easy-to-identify target species. Equivalent to around 20 species per Broad Habitat
Level 1	Record from a selection of c.200 easy-to-identify target species. Equivalent to around 10 species per Broad Habitat

Very few people would be able to undertake Level 3, but their contribution will be essential to provide surveillance of an unbiased sample of species from which unanticipated or unforeseen changes might be detected in the future. Seven hundred people participated in the BSBI's Local Change survey which records all species encountered and many more volunteers might be encouraged to participate at Level 3 if the survey method is straightforward.

Over 1000 people have participated at least once in Plantlife's Common Plants Survey. Experience has shown that the original list of species from which volunteers were asked to record (65 species) was too limited due to a combination of small sample size and plot sizes; for example, 32% of people did not encounter any of the target species in their plots. Therefore, in 2010 a larger sample of target species was decided on as a way to increase the likelihood of participants encountering species to record and consequently the list has been increased to 99 species. Level 1 would be important as a mechanism to significantly increase participation and the amount of data collected. It would encourage people to contribute to a worthwhile project and therefore improve their skills, with the potential for recorders to progress to Levels 2 or 3 eventually.

Level 2 could be seen as the core of the survey, appealing to the widest group of participants, including BSBI and Plantlife members, Breeding Bird Survey recorders, Wildlife Trust members, Butterfly Conservation recorders and other keen naturalists.

The following sections examine the methods available for the selection of target species for those participants preferring not to record all taxa.

3.2 Criteria for selecting target species

3.2.1 Habitat type

Stratification of target species by habitat type is a necessary feature of a plant surveillance scheme to ensure measurements are made of habitat quality which can be interpreted in terms of habitat-specific environmental drivers of change for plants and other wildlife. This method will target species information relevant to particular policy and practice delivery mechanism (e.g. agri-environment options, woodland grant schemes, open habitat policy).

Stratifying target species by habitat will ensure comprehensive representation of all UK habitat types because surveyors will need to identify all habitats in their sample plots. The method used to describe and identify habitat types should therefore be robust and relatively straightforward.

Several habitat classifications are currently in use in the UK. In the following subsection we assess the potential of each classification in providing a) the practicalities of volunteers recognising and recording them in the field; and (b) their relevance for analysis and reporting the results of a new scheme.

i. UK Broad Habitats

- This 'broad' habitat classification was initially developed as part of the UK Biodiversity Action Plan and later modified to take into account relationships to other British vegetation classifications (Jackson 2000; Hill, Preston & Roy 2004). Broad habitats are intended to be comprehensive and exclusive, so they should be easily interpreted on the ground. There are 21 Broad Habitat types that include vascular plants. The main advantages of this classification is that it is relatively easy for surveyors to apply in the field and that the species classification has been published and is widely available (PLANTATT; Hill *et al* 2004). The main disadvantage is that species are assigned to a maximum of four groups even if they are known to occur in more. Furthermore, some of the categories are very broad and contain many subtypes (e.g. Fen, Marsh and Swamp). In addition, some are not distinct habitats in an ecological sense (e.g. Bracken), describe habitat features rather than distinct communities (e.g. Boundary and Linear) and contain very few species (e.g. Bracken) or very few widespread species (e.g. Montane) such that both the habitat and target species it contains within it might not be encountered very frequently.

ii. National Vegetation Classification (NVC)

- The NVC is a detailed phytosociological classification of British vegetation types based on the component vascular plant, bryophyte and macro-lichen species. There are 286 recognised communities, organised into 12 major categories (PLANTATT provides a useful cross-reference between NVC communities and Broad Habitats). The main advantage is that it is relatively comprehensive allowing the majority of widespread species to be assigned to at least one NVC category, unlike many with other classification systems (e.g. Phase 1 habitats). As a consequence there is a relatively large pool of species (per category) from which to select a subset of targets, thereby increasing a sub-set target list can be selected. This increases the relative representativeness of the target list. The main disadvantage is that many ordinary communities, including anthropogenic types, are either not represented or poorly sampled. This is especially true of ecotones and mosaics which the NVC survey expressly avoided due to the complications that this would have created for the classification as a whole. Second, the recognition of NVC types requires a high-level of expertise and even then significant variation between surveyors is known to occur. Indeed most voluntary recorders or 'low-paid' professionals are unfamiliar with most if not all of the NVC types, often only having a 'working knowledge' of those present in their area.

iii. Phase 1 habitats

- The Phase 1 habitat classification is a very broad classification developed to help map vegetation types throughout the country. There are 155 specific habitat types in the classification divided into 10 (A-J) higher-level categories. The main advantages are its comprehensiveness and the ease by which it can be used by surveyors with minimal botanical or ecological expertise although few volunteers will be familiar with this system. Furthermore only 352 taxa (dominant species) are allocated to habitat types, which would limit the representativeness of a sub-set target species list.

iv. EUNIS Level II habitats

- The European Nature Information System (EUNIS) habitat classification is a pan-European system developed by the European Environment Agency (EEA) to cover all types of natural and artificial habitats, both aquatic and terrestrial. The classification is hierarchical, with 9 terrestrial habitats at Level I and 49 at Level II. The main advantages are its comprehensiveness which provides a level of detail between Broad Habitats and the NVC including additional 'habitats' subsumed within larger Broad Habitat categories (e.g. Fen, Marsh and Swamp). Its correspondence with both NVC and Broad Habitats are now widely available (Annex 2). The main disadvantage is that volunteers will be unfamiliar with it (although the classification is intuitive) as portions of the classification do not apply to GB.

Broad Habitats are used as the basis for assessments in the remainder of this report as they provide the most accessible and comprehensive classification already in use (Hill *et al* 2004). However, the consensus of workshop participants was that EUNIS Level II was more appropriate to use to aid the selection of a list of target species. The table in Annex 2 shows the correlations between Broad Habitats and EUNIS Level II habitats.

In terms of species selection, no habitats should be excluded but the plot selection method might mean that some habitats (such as montane) are not encountered frequently enough to achieve sufficient replication for statistical analyses. This situation might have to be addressed through stratification of sample locations or a targeted effort or additional project using professionals to improve the coverage of upland habitats.

3.2.2 Taxa included

At Level 3 recorders would attempt to record all species encountered including archaeophytes, neophytes, hybrids but not microspecies in the apomictic genera *Euphrasia*, *Hieracium*, *Rubus* and *Taraxacum*. Some recorders are likely to record infraspecific taxa and difficult segregates but many such taxa are likely to be treated as aggregates during analyses. The most recent edition of Clive Stace's (2010) flora is likely to be the accepted taxonomy that Level 3 volunteers are likely to follow for at least the next decade.

For Level 2 and 3 subsets of target species would need to be chosen using a range of criteria (see below). The initial pool of species would include around 1800 taxa for which information on plant traits are available in PLANTATT. These are all the taxa mapped in the text of the *New Atlas of the British and Irish Flora*, but not additional taxa included on the CD-Rom (Preston *et al* 2002). This pool of species includes all native species and some native subspecies, all archaeophytes plus around 250 commonly recorded neophytes (including all of the most invasive species) but excludes all microspecies of *Rubus fruticosus* and the large critical genera *Hieracium* and *Taraxacum* (included as single aggregates). The list also includes common hybrids and microspecies in some difficult genera (e.g. *Euphrasia*, *Limonium*, *Sorbus*) although few of these are likely to be selected as targets as they require expert determination.

3.2.3 Frequency of occurrence

Many species are too rare to be encountered sufficiently frequently during broad scale surveillance for meaningful change statistics to be derived for individual species. For example only 34% of species (860) recorded during the Local Change project were found in enough tetrads for statistical analyses to be performed (Braithwaite *et al* 2006)on the individual species. Selecting such rare and scarce species as targets is therefore undesirable as surveyors will become less inclined to participate if they do not encounter them. The size of the recording unit, as well as the sample size and sampling method, obviously is critical in determining detection rates within any surveillance scheme. This is illustrated in Figure 3.1 which shows the percentage occurrence of species recorded in hectad frequency classes for Countryside Survey, Common Plants Survey, Local Change and in the National Vegetation Classification.

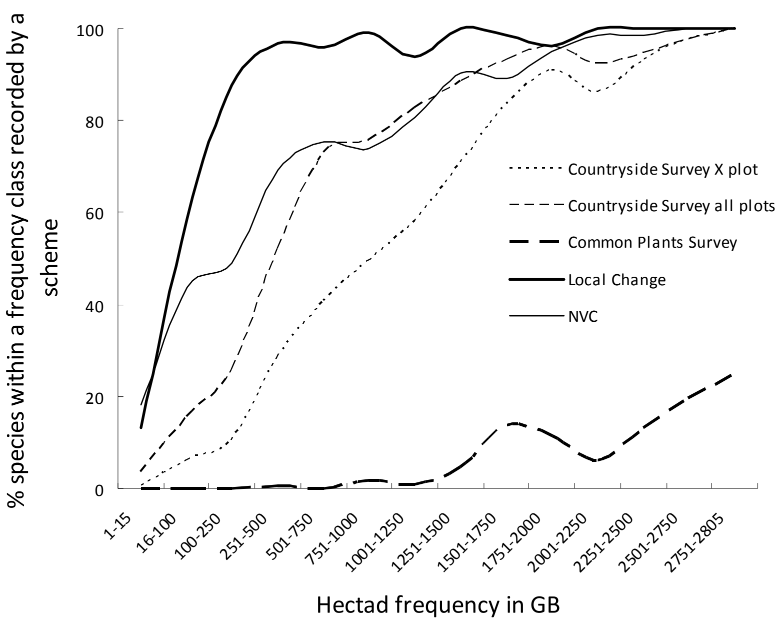


Figure 3.1. The percentage of GB species detected in hectad frequency classes for Countryside Survey, Common Plants Survey, Local Change and in the National Vegetation Classification plots. Note that the Common Plants Survey recorded 65 widespread species and would therefore not be expected to detect rare and scare taxa.

These schemes differ markedly in the size of the recording unit from entire tetrads (Local Change) to medium size plots (Countryside Survey, Common Plants Survey) to individual quadrats (NVC). As expected the results show that the larger the unit the more rare and scarce species are detected. Although small, I the NVC plots are also more likely to detect rarer species, presumably because they are focused on less common vegetation types that are more likely to support populations of ‘specialists’. However, the aim of designing a new surveillance scheme is not to record rare and scarce species, but detect changes in the wider countryside. Some rarer and scarcer taxa will be detected by recording Level 3 (recording all species encountered) but the selection of target species must focus on more widespread species.

Based on the detection rates in Local Change, NVC and Countryside Survey, it is proposed that only taxa found in more than 750 hectads are selected as target species. Whether to exclude very common species is debatable (e.g. the 50 most widespread species present in >2750 hectads in GB) and requires field testing. Such species are likely to occur in a high proportion of plots and are therefore time-consuming to record. Furthermore surveillance would only be able to detect negative trends for such species as they will be virtually ubiquitous across the sampling domain. However, changes in the abundance of such species (rather than occupancy) could provide very powerful indications of changes in habitat quality, especially given the large sample sizes that are likely to be available.

3.2.4 Ease of identification

To ensure that Level 2 and 1 surveyors can participate in the survey and that the results are statistically robust it is important that target species are easily identified and not confused with similar-looking species. A classification of British native and alien species according to the level of expertise needed for identification is currently being developed by the BSBI (Table 3.1).

Table 3.1. An example of a classification of British vascular plants according to level of expertise needed for identification purposes (source = BSBI)

	Level of determination needed				
Species	**National expert**	**County recorder**	**Other competent**	**Any recorder**	**Specimen needed**
Abies alba			x		
Abies grandis			x		
Abies procera			x		
Abutilon theophrasti		x	x		x
Acaena anserinifolia	x	x			x
Acaena anserinifolia x *inermis*	x				x
Acaena inermis	x	x			x
Acaena magellanica	x				x
Acaena novae-zelandiae		x			x
Acaena ovalifolia	x	x			x
Acer campestre				x	

This classifies species in relation to whether a determination is needed by a national expert or competent BSBI recorder such as a vice-county recorder (backed by a specimen that should be retained) or whether records can be accepted by any recorder regardless of level of expertise. This classification could be used as a basis to categorise species into three groups of difficulty (see below) of which only easy and moderately easy to identify species would be selected as target species.

Easy to identify	Species that are very straightforward to identify, with few similar species (excludes all grasses, sedges, rushes, most ferns)
Moderately easy to identify	Species requiring the use of keys and ability to differentiate between similar species (includes some grasses, sedges, rushes and ferns etc.)
Difficult to identify	Species requiring competent/expert determination

3.2.5 Indicator value (axiophytes)

Axiophytes are ‘worthy plants’, the 40% or so species that are of interest to botanists because they are usually wholly restricted to high quality habitats that are considered important in conservation terms (Lockton 2006). Although many axiophytes are rare species, rarity is not a defining quality *per se*. Indeed it could be argued that rare species are poor indicators because they occur on so few sites; in this sense they are likely to be ecologically (or statistically) trivial telling us little about the overall quality of most sites. In comparison, axiophytes that are much more widespread, i.e. those species that are always encountered in high quality sites, have much greater potential for indicating habitat quality in the wider countryside, and changes to populations of these species can tell us much more about the state of sites and factors driving change.

The selection of axiophytes first requires identifying habitats of conservation importance. Axiophytes are then species that are (a) 90% restricted to these conservation habitats and (b) recorded in fewer than 25% of tetrads in a county. An exception to the latter (25%) rule is made for species in conservation habitats that are particularly well represented and widespread in the county.

To date, 19 counties have produced lists of axiophytes (Cornwall, Dorset, South Hampshire, Sussex, Berkshire, Norfolk, Staffordshire, Shropshire, Cardiganshire, Lincolnshire, Cheshire, South Lancashire, Mid-west Yorkshire, Durham, Northumberland, Banffshire, Clyde Islands, Waterford and Antrim). Widespread species (i.e. occurring in greater than 750 hectads) that have been selected in more than 90% of the counties in which they occur are listed in Table 3.2. These cover a range of Broad Habitat types, in particular broad-leaved woodland and calcareous grassland. Remarkably three species have been selected as axiophytes in all 19 counties (*Carex laevigata*, *Sanicula europaea*, *Veronica scutellata*).

Axiophytes provide a useful category of species that could be used to monitor the ‘quality’ of habitats. It is very likely that some of the species finally chosen as targets will be axiophytes, and it would seem advantageous to use these to refine the final list once the criteria for their identification have been agreed and more counties have produced them.

Table 3.2. Widespread species (>750 hectads) listed as axiophytes in over 90% of the counties in which they are known to occur (*n* = 19, March 2010). An additional 138 species were recorded from less than 750 hectads. All are classified as needing to be determined by a competent recorder.

Taxon name	Number of GB hectads	Counties recorded	Counties axiophyte	% counties axiophyte	Level of ID difficulty
Berula erecta	1111	17	16	94	Comp
Carex laevigata	1012	19	19	100	Comp
Ceratocapnos claviculata	1122	18	18	100	Comp
Chrysosplenium alternifolium	790	14	14	100	Comp
Galium odoratum	1836	19	18	95	Comp
Genista tinctoria	932	15	15	100	Comp
Gentianella amarella	884	15	15	100	Comp
Helianthemum nummularium	1002	14	13	93	Comp
Juncus gerardii	919	16	15	94	Comp
Melampyrum pratense	1696	19	18	95	Comp
Ophioglossum vulgatum	1474	19	18	95	Comp
Orchis mascula	1962	19	18	95	Comp
Pimpinella saxifraga	1938	19	18	95	Comp
Polystichum aculeatum	1618	19	18	95	Comp
Ranunculus auricomus	1379	18	18	100	Comp
Sanguisorba officinalis	946	15	14	93	Comp
Sanicula europaea	2025	19	19	100	Comp
Silaum silaus	963	13	13	100	Comp
Vaccinium vitis-idaea	938	12	12	100	Comp
Veronica scutellata	1877	19	19	100	Comp
Viola reichenbachiana	1128	17	16	94	Comp

3.2.6 Specialist versus generalist species

Target species can be drawn either from habitat specialists (species that grow in one habitat type) or generalists (species that occur in several habitat types). Selecting specialists will naturally be more likely to reveal trends in particular habitats, such as the effect of shading in woodland and under-grazing in grasslands, which are more likely to be linked to drivers of change. Generalists are, however, more likely to highlight wider countryside issues and drivers of change, such as eutrophication, as they are less likely to be restricted to protected sites. The effect of restricting target species to habitat specialists on potential species selection is illustrated below. It is likely that it will be important to include some habitat specialists and some generalists in the target species pool. This approach would allow a comparison of trends between generalists and specialists, an approach that has delivered important results in analysing butterfly trends, and one which furthers the relevance of a plant surveillance scheme to measuring changes in the wider countryside.

One way to identify habitat specialists from existing data is to take those species restricted to a single Broad Habitat in PLANTATT. These are given in the final column of Table 3.3.

Table 3.3. Number of GB native and archaeophyte plants in relation to broad habitats and frequency of occurrence (number of hectads in GB). An asterisk denotes where there are likely to be insufficient species within a Broad Habitat to allow selection of sufficient targets to allow meaningful analyses (See 3.3 below).

		Frequency of occurrence (number of hectads in GB)					
Broad Habitat name	**BH code**	**<750**	**751-1500**	**1501-2250**	**2251-2805**	**>750**	**1 BH**
Broadleaved woodland	BH1	108	38	50	45	133	52
Conifer woodland	BH2*	11	2	4	4	10	0
Boundary & linear	BH3	142	77	53	60	190	38
Arable	BH4	54	30	33	21	84	24
Improved grassland	BH5*	5	2	0	9	11	1
Neutral grassland	BH6	45	29	19	52	100	32
Calcareous grassland	BH7	111	46	23	21	90	35
Acid grassland	BH8	43	13	10	13	36	10
Bracken	BH9*	1	1	1	7	9	0
Heath & mire	BH10	35	14	11	6	31	5
Fen, marsh & swamp	BH11	97	55	45	39	139	64
Bog	BH12	20	4	14	1	19	7
Standing water & canals	BH13	90	33	15	3	51	4
Rivers & streams	BH14	44	30	23	15	68	5
Montane	BH15*	86	8	0	0	8	0
Inland rock	BH16	155	40	29	19	88	7
Urban	BH17	24	24	24	15	63	1
Coastal habitats	BH18-21	139	19	7	4	30	11

3.2.7 Plant traits

Information is available for a wide range of plant traits which could then be used to select a target species pool (Hill *et al* 2004). Those most useful for the purposes of surveillance are likely to be (1) Broad Habitats and frequency; (2) Ellenberg indicator values for light (L), moisture (F), reaction (R), nitrogen (N) and salt tolerance (S); (3) maximum summer plant height; (4) mean January and July temperatures; and (5) Raunkiaer life form. For each trait, target species could either be drawn from across the range of values (in proportion to the number of species in each category), from either extreme, or from functional groups that combine a combination of traits.

3.3 Using these criteria to select Broadleaved Woodland targets

In this section we illustrate how the above criteria might be used in the selection of target species for Broadleaved Woodland. Note that a strong recommendation from the surveillance workshop was the adoption of a EUNIS Level II classification for British vascular plants for both the initial selection of target species and for defining habitat types in the field (see section 3.2.1). This would follow the habitat classification of bryophytes in Britain and Ireland (Hill *et al* 2007) and align any new scheme with the widespread use of this classification in Europe. The translation of the species classification from Broad Habitats to EUNIS Level II is still to be completed. In the meantime, the following sections are based on the current classification of species by Broad Habitat. This is purely illustrative, however, and this work will need to be repeated once all species have been assigned to the EUNIS Level II classification.

The selection of target species needs to be stratified by habitat type to ensure sufficient surveillance coverage across all types. Table 3.3 quantifies the number of taxa in each Broad Habitat in relation to frequency of occurrence at the GB level. Note that the application of the widespread frequency criteria (>750 hectads) removes virtually all the species in the Conifer Woodland, Improved Grassland, Bracken and Montane Broad Habitats (denoted by an asterisk in Table 3.3). It would therefore seem sensible to reduce the threshold for these habitats in any future selection process.

3.3.1 Selection by frequency of occurrence

Within each broad habitat it would seem sensible to select target species based on frequency of occurrence (Table 3.3). Under this method several species could be selected from each frequency category. An illustration is given in Table 3.4 where we have selected two easily identifiable woodland species from 8 frequency categories (steps of 250 hectads from 750 onwards).

Table 3.4. Example of target species list for Broadleaved Woodland (BH1) based on selection of two species from each frequency category (number of hectads in GB in bold). Codes for traits follow Hill *et al* (2004) and are: **Chg**, relative Change Index (Preston *et al* 2002); **Hght**, maximum summer plant height (cm); **LF1**, primary Raunkaier lifeform (Gb = bulbous geophyte, Gn = non-bulbous geophyte, hc = hemicryptophyte, Ch = chamaeophyte, Pn, nano-phanerophyte, Ph = phanerophyte); **GB**, number of hectads in GB; **Tjan** & **Tjul**, mean January and July temperatures for hectads in which a taxon has been recorded; **L**, **F**, **R**, **N**, **S**, Ellenberg indicator values for light, moisture, reaction, nitrogen and salinity respectively.

Species	Chg	Hght	LF1	GB	Tjan	Tjul	L	F	R	N	S
Lonicera periclymenum	-0.11	600	Ph	**2622**	3.6	14.5	5	6	5	5	0
Silene dioica	-0.44	90	hc	**2514**	3.4	14.6	5	6	6	7	0
Ajuga reptans	-0.56	30	hc	**2439**	3.4	14.6	5	7	5	5	0
Stellaria holostea	-0.56	60	Ch	**2372**	3.5	14.7	5	5	6	6	0
Mercurialis perennis	-0.65	40	hc	**2214**	3.2	14.8	3	6	7	7	0
Allium ursinum	0.24	45	Gb	**2034**	3.5	14.8	4	6	7	7	0
Viburnum opulus	-0.15	400	Ph	**1854**	3.6	15.0	6	7	6	6	0
Galium odoratum	-0.62	45	hc	**1836**	3.4	14.7	3	5	7	6	0
Arum maculatum	-0.28	50	Gn	**1604**	3.9	15.3	4	5	7	7	0
Melica uniflora	-0.04	60	hc	**1511**	3.5	15.0	4	5	7	5	0
Carpinus betulus	0.84	3000	Ph	**1488**	3.6	15.4	4	5	5	6	0
Euonymus europaeus	0.15	600	Ph	**1254**	3.9	15.4	5	5	8	5	0
Lamiastrum galeobdolon	1.07	60	Ch	**1097**	3.7	15.8	4	5	7	6	0
Prunus padus	0.58	1500	Ph	**1089**	2.7	13.8	5	6	6	7	0
Campanula latifolia	-0.23	120	hc	**944**	2.9	14.6	4	5	7	6	0
Daphne laureola	0.10	100	Pn	**844**	3.7	16.0	4	5	7	5	0

3.3.2 Selection based on trait groups

The target species pool could be drawn proportionately from categories within a particular trait (Figure 3.2). In this way, all trait categories would be included providing a more representative sample of species for a particular trait.

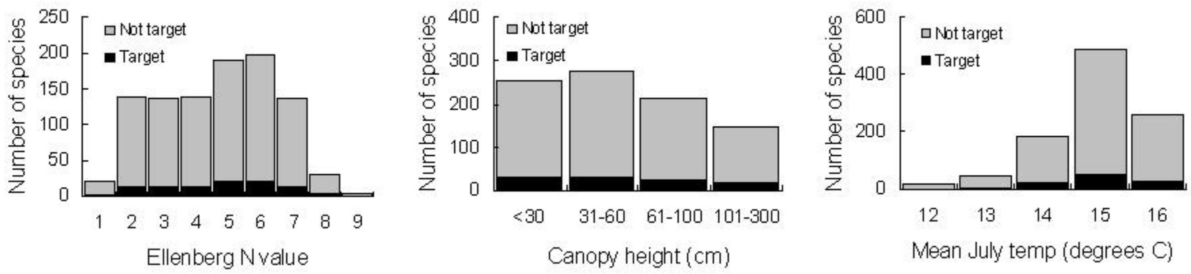


Figure 3.2. Selection of target species (black) in proportion to the total number of species in each trait category.

Table 3.5. Selection of Broadleaved Woodland (BH1) species based on proportion of species in (A) Ellenberg light (L) and (B) nitrogen (N) categories (highlighted in bold). Codes for traits are the same as in Table 3.4.

Species	Chg	Hght	LF1	GB	Tjan	Tjul	L	F	R	N	S
(A) Selected in relation to Ellenberg L											
Salix caprea	0.34	1000	Ph	2412	3.4	14.6	**7**	7	7	7	0
Rubus idaeus	-0.09	150	Pn	2425	3.4	14.5	**6**	5	5	5	0
Viburnum opulus	-0.15	400	Ph	1854	3.6	15.0	**6**	7	6	6	0
Ranunculus auricomus	-0.33	40	hc	1379	3.3	15.1	**6**	7	6	5	0
Cardamine flexuosa	1.06	50	hc	2580	3.5	14.5	**5**	7	6	6	0
Silene dioica	-0.44	90	hc	2514	3.4	14.6	**5**	6	6	7	0
Ajuga reptans	-0.56	30	hc	2439	3.4	14.6	**5**	7	5	5	0
Allium ursinum	0.24	45	Gb	2034	3.5	14.8	**4**	6	7	7	0
Arum maculatum	-0.28	50	Gn	1604	3.9	15.3	**4**	5	7	7	0
Galium odoratum	-0.62	45	hc	1836	3.4	14.7	**3**	5	7	6	0
(B) Selected in relation to Ellenberg N											
Elymus caninus	0.27	110	hc	1669	3.3	14.9	7	6	7	**8**	0
Allium ursinum	0.24	45	Gb	2034	3.5	14.8	4	6	7	**7**	0
Viburnum opulus	-0.15	400	Ph	1854	3.6	15.0	6	7	6	**6**	0
Galium odoratum	-0.62	45	hc	1836	3.4	14.7	3	5	7	**6**	0
Ranunculus auricomus	-0.33	40	hc	1379	3.3	15.1	6	7	6	**5**	0
Ajuga reptans	-0.56	30	hc	2439	3.4	14.6	5	7	5	**5**	0
Adoxa moschatellina	-0.05	12	Gn	1720	3.3	14.9	4	5	6	**5**	0
Epipactis helleborine	0.08	80	Gn	1218	3.6	15.3	4	5	7	**4**	0
Polypodium vulgare	-0.03	40	hc	2496	3.5	14.4	5	5	5	**3**	0
Platanthera bifolia	-1.67	40	Gn	949	3.5	14.3	6	6	6	**2**	0

Table 3.5 gives a selection of Broadleaved Woodland species selected in this way based on the proportion of species across Ellenberg L and N scores.

Alternatively, species could be selected from either end of their trait classes (selection of extreme trait groups). In this way, only the species exhibiting extreme traits for the habitat would be selected (Figure 3.3).

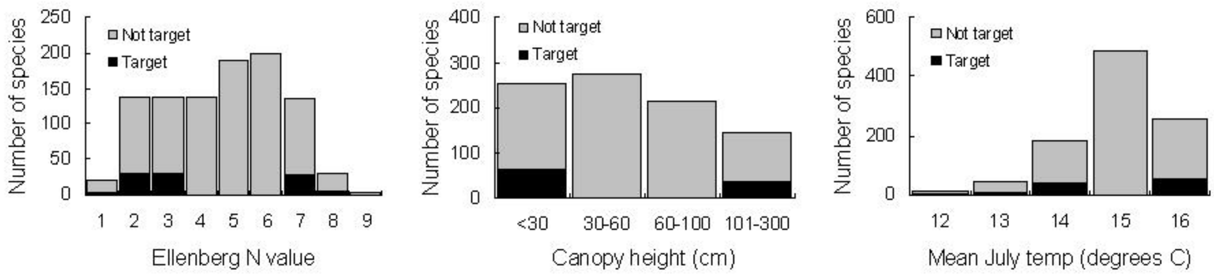


Figure 3.3. Selection of target species (black) from extreme ends of each trait category.

Table 3.6 gives a selection of Broadleaved Woodland species selected in this way again based on the proportion of species across Ellenberg L and N scores.

Table 3.6. Selection of Broadleaved Woodland (BH1) species from extreme trait groups for Ellenberg light (L) (highlighted in bold). Codes for traits are the same as in Table 3.4.

Species	Chg	Hght	LF1	GB	Tjan	Tjul	L	F	R	N	S
(A) Selected in relation to Ellenberg L											
Salix caprea	0.34	1000	Ph	2412	3.4	14.6	**7**	7	7	7	0
Rubus idaeus	-0.09	150	Pn	2425	3.4	14.5	**6**	5	5	5	0
Brachypodium sylvaticum	-0.17	95	hc	2310	3.7	14.7	**6**	5	6	5	0
Viburnum opulus	-0.15	400	Ph	1854	3.6	15.0	**6**	7	6	6	0
Ranunculus auricomus	-0.33	40	hc	1379	3.3	15.1	**6**	7	6	5	0
Allium ursinum	0.24	45	Gb	2034	3.5	14.8	**4**	6	7	7	0
Arum maculatum	-0.28	50	Gn	1604	3.9	15.3	**4**	5	7	7	0
Campanula latifolia	-0.23	120	hc	944	2.9	14.6	**4**	5	7	6	0
Mercurialis perennis	-0.65	40	hc	2214	3.2	14.8	**3**	6	7	7	0
Galium odoratum	-0.62	45	hc	1836	3.4	14.7	**3**	5	7	6	0
(B) Selected in relation to Ellenberg N											
Elymus caninus	0.27	110	hc	1669	3.3	14.9	7	6	7	**8**	0
Mercurialis perennis	-0.65	40	hc	2214	3.2	14.8	3	6	7	**7**	0
Silene dioica	-0.44	90	hc	2514	3.4	14.6	5	6	6	**7**	0
Arum maculatum	-0.28	50	Gn	1604	3.9	15.3	4	5	7	**7**	0
Allium ursinum	0.24	45	Gb	2034	3.5	14.8	4	6	7	**7**	0
Betula pendula	-0.23	2500	Ph	2293	3.3	14.6	7	5	4	**4**	0
Quercus petraea	0.14	3000	Ph	1832	3.5	14.6	6	6	3	**4**	0
Dryopteris carthusiana	1.06	80	hc	1623	3.3	14.6	6	8	5	**4**	0
Polypodium vulgare	-0.03	40	hc	2496	3.5	14.4	5	5	5	**3**	0
Platanthera bifolia	-1.67	40	Gn	949	3.5	14.3	6	6	6	**2**	0

Alternatively, species could be classified into combined trait groups based on up to three traits. Species would then be selected from each grouping. The example in Table 3.7 uses July temp, soil fertility (Ellenberg N) and plant (canopy) height and shows the number of species in each category from which target species could be selected.

Table 3.7. Number of Broadleaved Woodland (BH1) species available for selection using a combination of three traits; July temperature, soil fertility (N) and canopy height.

July temp →	**Low**			**Intermediate**			**High**		
Soil fertility →	**Low**	**Intermediate**	**High**	**Low**	**Intermediate**	**High**	**Low**	**Intermediate**	**High**
Canopy height ↓									
<0 (aquatics)	10	5	0	0	20	3	0	6	7
1-30 cm	75	23	1	41	57	13	25	17	2
31-100 cm	68	34	2	35	158	54	25	88	31
101-300 cm	7	12	1	6	43	30	3	21	15
>300 cm	1	4	1	0	26	5	0	13	5
Total	161	78	5	82	304	105	53	145	60

Although the use of traits to select targets would seem desirable the number of potential traits that could be used would make the selection process overly complex. We therefore recommend that traits are excluded from the initial process but are used to check for representativeness and bias following selection using other attributes (e.g. broad habitat, frequency of occurrence, ease of identification, etc).

3.3.3 Selection based on ease of identification

The selection of target species in each habitat could be based on their ease of identification, with only species requiring low to moderate botanical expertise being sampled. However, this could bias the data by selecting an unrepresentative sample in relation to certain plant traits, geographic distributions, etc. In order to assess this potential effect we examined the means and standard deviations of different subsets of species based on ease of identification. In doing this we restricted the sample to widespread species (>750 hectads) restricted to woodland (n = 52). The means for various traits were then calculated for three subsets: (1) easy to identify, (2) easy and moderately easy to identify and (3), easy, moderate and difficult species to identify (Figure 3.4).

The figures indicate that, by sampling the easy to identify species only, we could just about expect to sample averages within the same standard deviation as the full sample, but levels of variability are more robust when sampling from easy and moderate group of 43 species.

3.3.4 Selection based on a random draw

A problem with the selection of species described in 3.3.3 is the tendency to select an unrepresentative selection of species. To overcome this, we could select species from each Broad Habitat in an entirely random way. Figure 3.5 shows the effect of sampling 10, 20, 30, 40, 50 and 60 species randomly from the total sample of 136 widespread woodland species (recorded in >750 hectads) on mean values for various traits.

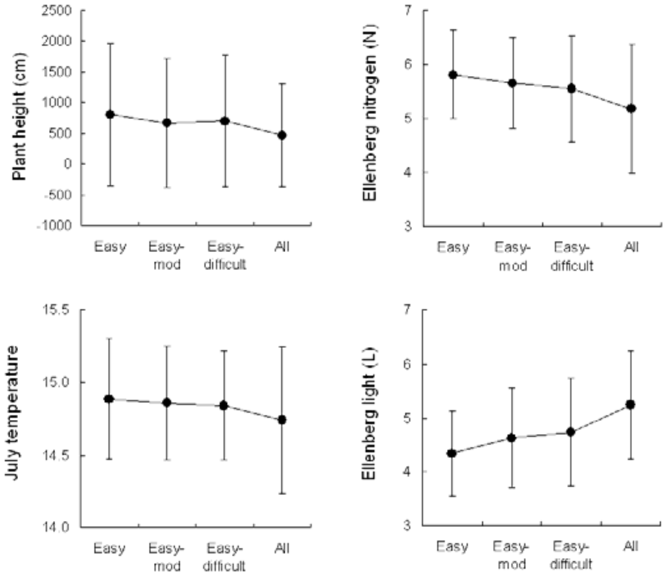


Figure 3.4. Effect of sampling specialist broadleaved woodland (BH1) target species, based on their ease of identification, in relation to various traits (plant height, July temperature, soil fertility N and light requirements L). The data are means ± 1 standard deviation. Three subsets of species are shown: easy to identify species (n = 21), easy and moderately easy to identify species (n = 43) and easy, moderate and difficult species to identify (n = 52). These are compared to the mean of all 136 species classified as woodland species regardless of whether they occur in other habitats.

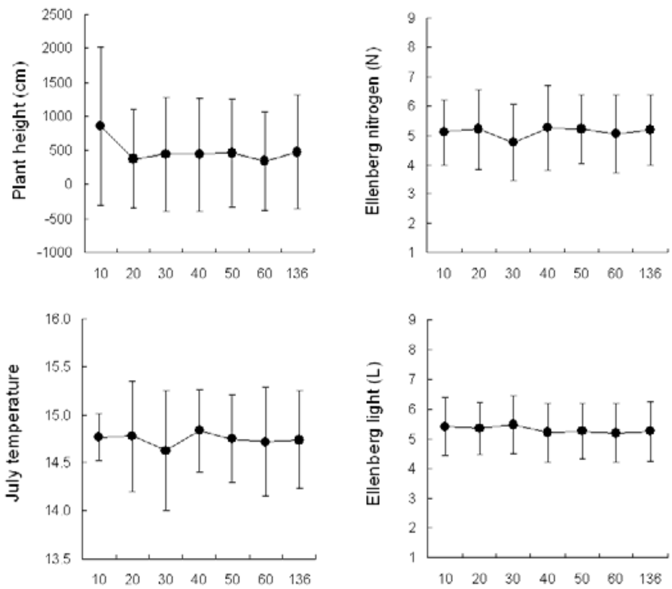


Figure 3.5. The effect of sampling 10, 20, 30, 40, 50 and 60 species chosen randomly (single randomisation trial only) from the total pool of 136 woodland species that are recorded from >750 hectads on the mean (squares with standard deviation), values of various traits (height, July temperature, soil nutrients N and light requirements L).

The results show that the variation in traits can be captured within relatively few species, and a random draw can provide a reasonable first selection.

3.4 Conclusions on species selection

Several methods for selecting target species have been examined. Following feedback from discussion with project partners and at the workshop, and the assessments described above it is recommended that target species are not selected on the basis of traits. This would inevitably lead to bias in the sample by selecting for traits related to known drivers of change (soil fertility for example), reducing the effectiveness of the survey to detect changes in all drivers and unknown drivers in the future that cannot be predicted. It will also be difficult to select the most appropriate group of species representing all traits across all habitat types. Having said that an important goal of the species selection should be to ensure that targets are representative of the full range of traits, rather than extremes or key species. They should also be easy or moderately-easy to identify, to ensure wide participation, with 10-30 are available for each habitat type at levels 1 and 2 to ensure sufficient replication for reporting trends at the habitat level (actual number per habitat may differ with use of EUNIS categories since there are a greater number of these than Broad Habitats). Workshop participants also felt it was important to include species that are important food resources for other taxa (e.g. pollinators). The potential confounding effect of phylogenetic relatedness (autocorrelation) was also raised as closely related indicator species will tend to behave more similarly in comparison to more distantly related species. If possible it would therefore seem sensible to avoid selecting targets from the same genera or families (although the latter may be unavoidable).

In conclusion, it has not proved possible to agree a single predefined set of target species, However, restricting the species recorded to those that are easier to identify does not cause bias in a range of traits. An approach to agreeing a list of target species is given:

1. Draw up an initial pool of species that would be used in analyses (all native with exception of some apomictic groups).
    - *This is the **Level 3 list.***
2. Stratify this pool by habitat noting that :
    - Some species will appear in multiple strata;
    - This information can then be used to derive a specialist/generalist classification.
3. Restrict the lists to those species occurring in >750 hectads (and possibly also <2750 hectads but this requires field testing).
4. Restrict the list to those that are easy to moderate to identify.
5. Produce a random draw of 10-20 species from each habitat (subject to modification for use of EUNIS II categories).
6. Produce an additional random draw of a further 10 species from each habitat.
7. Check representivity of a range of traits (focusing on specialists/generalists, pollinators and food plants) in total list of species selected for all habitats:
    - *These are the **Level 2** and **Level 1 lists.***

We propose that participants are encouraged to record either all species they encounter on the Level 3 list, if they are sufficiently experienced to do so, or from Level 2 and 1 lists depending on experience and skills. This will significantly improve the level of participation, increase the number of sites surveyed, and ultimately improve the quality of the data collected.

In addition the following recommendations are considered high priority for the development of a new scheme:

- Target species should be stratified according to habitat and based on a EUNIS Level II classification. Development of this classification should be given priority
- Within each habitat, a variety of both specialist and habitat generalist species should be identified. If possible the selection procedure should follow those used for birds and butterflies to ensure a common approach between taxon groups. Plant specialists might be defined as species restricted to a single EUNIS Level II habitat and/or widespread axiophytes. Conversely generalists will be common species that occur in a range of habitat types.

4 Sample stratification

4.1 Sample unit

Kilometre grid squares (1 × 1 km), otherwise termed monads, provide a useful sampling unit in which to record vascular plants. In recording terms they provide a manageable 'area' for volunteers to survey in a single day, depending on terrain and access, even within heterogeneous lowland landscapes. Most contain a range of habitats, even in apparently uniform terrain, such as upland bog, where small habitat features, such as flushes, pools, mires, add significantly to overall diversity. Consequently, they have been used to stratify surveillance sampling for plants (Common Plant Survey and Countryside Survey) as well as a range of other taxonomic groups, including birds and butterflies.

Monads are increasingly being used by volunteers as the basis for geographic sampling at the county scale, especially in southern England where more volunteers are available to undertake recording smaller units. In comparison, larger units such as tetrads (2 × 2km), pentads (5 × 5 km) and hectads (10 × 10 km) are less 'manageable' in the sense that repeated visits are needed to cover the sample area due to greater habitat and species diversity.

The number of species likely to be encountered in an 'average' monad can be gauged from the average number of taxa recorded in tetrads monitored as part of the BSBI Monitoring Scheme (Figure 4.1). These were first recorded in 1986/7 and revisited in 2003/4 (Braithwaite *et al* 2006). The mean number of taxa for both surveys were 216 (±3) and 259 (± 4) respectively. We would therefore expect to record 200-300 species in an 'average' monad, although the exact number is uncertain due to the change in recording scale.

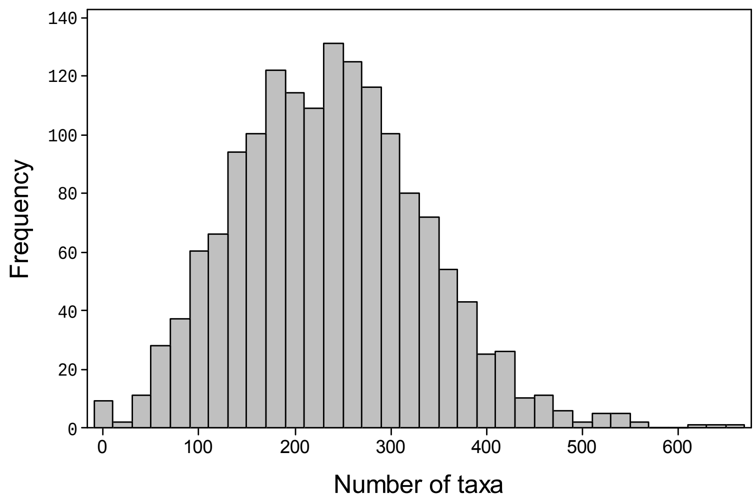


Figure 4.1. The number of plant taxa recorded in Monitoring Scheme tetrads in 1987/88 (n = 808) and 2003/04 (n = 760).

4.2 Stratification

Geographic stratification of sampling is required within monitoring or surveillance schemes to ensure sufficient replication across the full range of environmental gradients within the sampling domain. Consequently an 'ideal' scheme would aim to sample within each geographic unit however defined. For vascular plants the most widely recognised ‘recording strata’ are Watsonian vice-counties as these have formed the basis of voluntary recording activity since they were devised over 100 years ago (Figure 4.2). The vice-counties were introduced by Hewett Cottrell Watson in 1852 and with the exception of minor amendments they have remained unchanged since, allowing historical and modern data to be more accurately compared. Other alternatives include environmentally defined strata, such as the Environmental Zones used in Countryside survey (Figure 4.2). Although voluntary recorders are likely to be unfamiliar with them this need not be a problem as recorders could just be sent the ‘squares’ in their county following stratification. However, they are less practical if the densities of recorders in vice-counties are used to help define the number of strata in each area as it would be difficult to relate these area figures to EZs (see below).

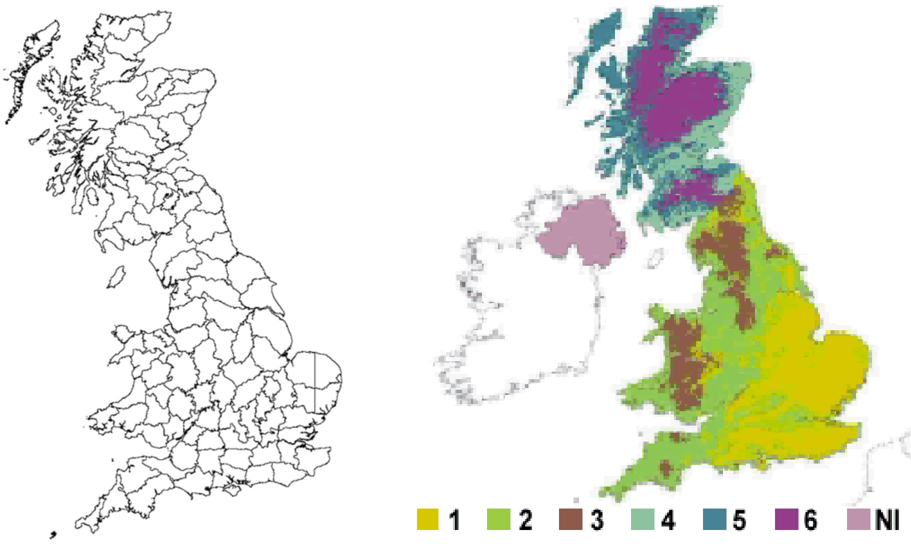


Figure 4.2. Two approaches to geographic stratification: Watsonian vice-counties (left) and Environmental Zones (right) used to stratify sampling in Countryside Survey (1, Easterly lowlands (England/Wales); 2, Westerly lowlands (England/Wales); 3, Uplands (England/Wales); 4, Lowlands (Scotland); 5, Intermediate uplands and islands (Scotland); 6, True uplands (Scotland); 7, Northern Ireland).

4.3 Sampling within strata

The selection of units for sampling within geographic strata could include random, systematic or *ad hoc* approaches. Random sampling would be the ideal approach with the sample for each county generated prior to survey. These would then be allocated to surveyors on a ‘first-come, first-served’ basis (as currently occurs in the Breeding Bird Survey) with ‘new’ random squares generated only when all available random squares had been allocated. Alternatively, squares could be randomly selected within a predetermined distance (say 5 or 10km) of a volunteer’s home, as is done with the Common Plants Survey.

In comparison, a systematic sample could be generated, most likely based on a fixed grid of monads superimposed over the vice-county system. This is identical to the approach currently used in Local Change where three fixed tetrads (‘A, J & W’ using the DINTY naming convention) are recorded on a fixed grid of 1 in 9 hectads (Braithwaite *et al* 2006). The distribution of these fixed tetrads is shown on Figure 4.3a. Such a systematic approach poses a number of challenges not least the difficulty of recording in sparsely populated areas and/or under-sampling of restricted habitats, such as mountains, rivers, coastline, that are poorly represented on a fixed grid. A separate survey effort or programme could be employed (such as a montane recording scheme) to improve recording of such squares.

The least desirable of all options, from a statistical view point (but probably most favoured amongst recorders), would be *ad hoc* recording where volunteers 'select' squares for sampling within their own vice-county. This would inevitably lead to a bias towards 'home squares' or richer, more varied squares containing a higher diversity and an under-sampling of more uniform, less diverse countryside.

Rather than selecting a completely new set of squares, a more pragmatic alternative would be to select a random sample of squares drawn from an existing 'pool' of squares already being recorded for other surveillance schemes. The most relevant are Local Change, Common Plants Survey, BTO Breeding Bird Survey and BC Wider Countryside Butterfly Scheme (Table 4.1). The distributions of squares for three of these schemes are shown in Figure 4.3.

Table 4.1. Existing surveillance schemes relevant to the development of an 'ideal' scheme for plants.

Surveillance scheme	Sample size (km)	Stratified sample?	Random sample?	Details
BSBI Local Change (LC)	2	No	No	Systematic sample of tetrads in 1 in 9 hectads
Plantlife Common Plant Survey (CPS)	1	No	Yes	Random allocation of 1km squares from 10km square grid around surveyors home postcode
BTO Breeding Bird Survey (BBS)	1	Yes	Yes	Stratified-random sample, stratified by 80 BTO regions (comparable to counties) and weighted by number of BTO recorders per region
BC Wider Countryside Butterfly Scheme (WCBS)	1	Yes	Yes	Stratified-random sample, combination of existing BBS squares and squares covered by BC members (stratified-random sample, stratified by BC branches and weighted by number of BC recorders per region)

Figure 4.3. Distribution of sample squares currently monitored as part of the following surveillance schemes: Local Change tetrads (top left), Common Plant Survey monads (top right), Wider Countryside Butterfly Scheme monads (bottom left) and all schemes combined (bottom right).

Such an approach has a number of advantages. Firstly, practical experiences of surveying squares, including gaining access permissions, which is known to be one of the main limiting factors to surveys, could be shared between schemes. Second, 'sharing' squares in this way would allow multi-taxon comparisons to be made, thereby facilitating improved interpretation of trends. Third, by randomly selecting from squares already under surveillance, non-random squares, such as those included in Local Change, could be incorporated into the sample, thereby improving the baseline from which trends can be drawn. Finally, as with the BTO and BC schemes, such an approach might encourage participation by surveyors from different taxonomic schemes.

4.4 Number of samples within strata

The number of samples within strata is irrelevant where using a systematic grid-based approach but would require some form of weighting if strata were geographically defined. Ideally this would be proportional to the area of the strata to ensure even coverage of recording. However, this poses a number of challenges, especially for recording in areas with few recorders that are also often difficult to record because of inaccessibility and/or terrain (e.g. uplands, islands, etc.). This is shown clearly in Figure 4.4 which displays the current distribution of BSBI members at the hectad scale. The lack of recorders 'on the ground' is particularly striking in the highlands and islands of Scotland, the Borders, some parts of northern and eastern England, the uplands of Wales and large parts of Northern Ireland (no data for the Republic).

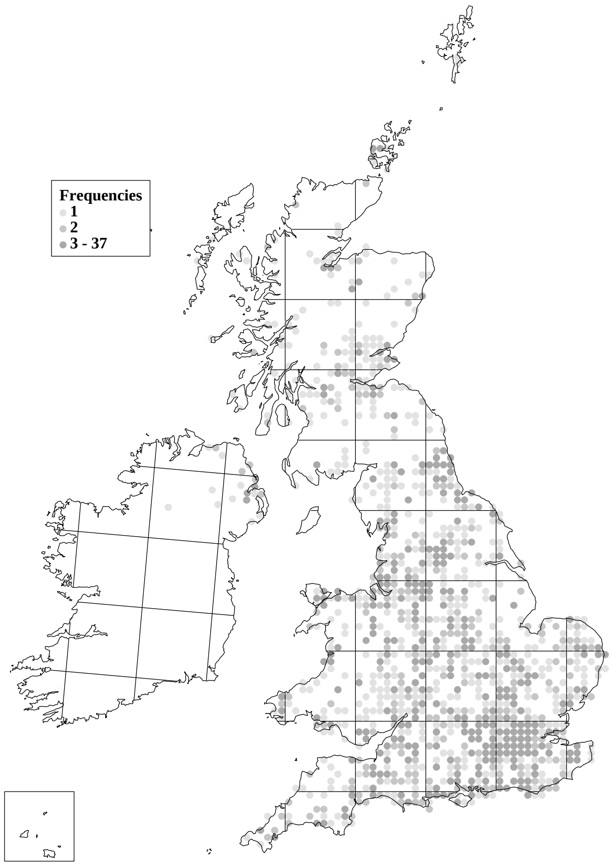


Figure 4.4. Coincidence map of the density of BSBI members at the hectad scale in Britain and Northern Ireland (data for the Republic of Ireland are omitted).

A more pragmatic solution, as used in other surveillance schemes (Table 4.1), is to concentrate sampling in more populous areas by weighting the number of samples by the number of recorders present in each strata, using members as a proxy for human population density (as shown for BSBI members in Figure 4.4). Obviously this presents a bias in sampling towards more populous regions but this would be minimized by setting a minimum and maximum of samples that can be recorded in any one region.

4.5 Conclusions on sample stratification

We propose a stratified-random sample of 1km squares as the basis of an ideal plant surveillance scheme. Strata would be Watsonian vice-counties and the number of samples per strata would be weighted to the number of recorders present, with a minimum and maximum number to reduce geographic bias in the sample. The selection of samples from within strata would be randomly selected from the ‘pool’ of random 1km squares in each county currently being used in existing schemes, including other taxonomic groups, plus the non-random squares included in Local Change. For the latter we would suggest using the monad in the SW corner of the tetrad, although a selection of any of the other monads would be equally valid. It is recognised that should some habitats be under-represented then supplementary surveys using skilled volunteers or professionals may be required.

5 Survey method

5.1 Introduction

In this chapter we assess a range of survey methods that could potentially be used to monitor trends in the status of plant species within a national sample of 1km squares. The monad should therefore be considered the site or 'primary sampling unit' for any new scheme. Recording all species or habitats in the entire monad would be impossible given the time and resources available. In this chapter we therefore consider the potential components for sampling the monad, including the measures needed for analyses, plot types and field methods for combining the surveillance components. The statistics that might be derived from each method and its policy relevance is given in the final section.

It should be noted that decisions on the overall design of the field survey have been made in the context of the survey criteria set out in Chapter 1 namely:

- A simple method achievable in 0.5-1 day survey per sample (depending on accessibility and distance to travel);
- Low cost, achievable through volunteers or restricted use of professionals;
- Attractive and engaging to volunteers;
- Ability to report status and trends for species/habitats as well as local information;
- Complementary to other surveillance schemes.

5.2 Recording abundance or frequency?

Unlike other taxonomic groups the estimation of the abundance (i.e. the number of individuals) of higher plants is almost impossible at scales of more than a few metres. Likewise a simple measure of presence/absence from a plot has very low power when it comes to statistical analyses (see Chapter 8). Therefore, a measure of frequency within a primary sample is commonly used in botanical recording as a proxy for abundance, and at a range of scales from tens of metres to hectads. Such approaches are more appropriate for volunteers because of their simplicity, as they only require a basic knowledge of ecological survey techniques, and a relatively small amount of time needed to survey the sample area. Consequently more volunteers are likely to take part in any scheme utilising measures of occurrence rather than true abundance. We have therefore focussed on approaches to recording frequency using a variety of selection methods and plot types. A single method for measuring abundance (i.e. percentage cover) is also included, but as a 'nest' within samples recorded under Methods 2-5 in this chapter.

5.3 Potential plot types

In the following section we present a range of survey methods based on sampling within large, small and nested plots and transects.

5.3.1 Large (hectare) plots

Large (hectare) plots based on the national OS grid (i.e. 100 × 100m grid cells) provide a convenient unit for recording the occurrence of higher plants for a number of reasons. First, most voluntary recorders use these routinely to record the distribution of rare or interesting species. Second, they provide a ‘manageable’ area that can be surveyed exhaustively in a relatively short period of time, even in heterogeneous landscapes or mountainous terrain. Third, they can be easily relocated using OS maps or handheld GPS and, because of their size, they are unlikely to be affected by changes in the positions of the boundary during resurveys. Fourth, most cells are likely to be dominated by a few major habitats, although linear features and small habitats features, such as ponds, flushes, etc. would obviously add to overall diversity.

The main disadvantage of large plots is potentially the high diversity of species and habitats present and therefore the time required to survey them exhaustively. This is only a problem though for experienced botanists recording all species encountered; for volunteers recording from a list of target species it is less of a problem, although recorders not used to 100m grid cells may find the size daunting. Small features that add significantly to overall diversity might also be easily overlooked at this scale, especially in heterogeneous landscapes. Also relocation may be unreliable where recorders do not use GPS and/or are unfamiliar with grid-based recording.

Despite these caveats, hectare plots provide a practical compromise between the rigour needed to record small plots and the desire of most recorders to survey as great a range of habitats and species present as possible.

5.3.2 Small plots

Small plots are generally taken to mean plots of dimensions less than 10 × 10m, although this obviously varies depending on the habitat type being surveyed. Such plots, often termed quadrats (or relevés), are mainly used in phytosociological studies to describe vegetation types based on their overall species composition, frequency and abundance. They are also commonly used to record the assemblages of rare or threatened species as well as to monitor changes in species composition, often in response to a changing explanatory variable, such as management or an experimental treatment. Quadrats can either be permanent or randomly located, although for the latter much larger samples sizes are required to detect change.

The key difference from larger plots is the attempt to provide a complete census of all species present, regardless of flowering stage, and to assign a measure of abundance. However, even at this small scale a complete count of individuals is usually not possible for the majority of species and so estimates are made based on a visual estimate of cover using conventional scales of abundance (e.g. DOMIN, DAFOR). However, this is not a true reflection of abundance as a grass covering half the quadrat may have far fewer individuals than a tiny annual that is sparsely distributed throughout the same quadrat but with an overall cover of less than 1%.

Despite these caveats, the main advantage of this approach is that it allows changes in abundance to be monitored very precisely, often directly to changes in explanatory variables. However, the major disadvantage for a voluntary scheme is their labour intensive nature. Quadrats can be very time consuming to record, especially where large numbers of replicates are needed to sample a community comprehensively. The techniques used also require a degree of training. Consequently, they are only likely to appeal to a tiny minority of voluntary recorders.

5.3.3 Nested plots

Nested plots have long been used by professional ecologists to monitor changes in vegetation. Various designs have been proposed (e.g. Stohlgren *et al* 1995; Critchley & Poulton 1998) but all have the same underlying principle; to record the first occurrence of species within plots of increasing size 'nested' in larger plots (Figure 5.1). At small scales the overall assumption is that the smaller the cell in which a species occurs first, the more abundant it will be, all other things being equal. Therefore, the rank order of occurrence can be used as an objective proxy for abundance or frequency.

The two main advantages of nested plots over standard quadrats are firstly, the more objective derivation of abundance measures and second, the more efficient searching of the sample area. Intensive searching of the smaller nests leads to a rapid accumulation of more widespread species. Subsequent searching of larger nests then focuses on the 'discovery' of new species, which by definition are more localised.

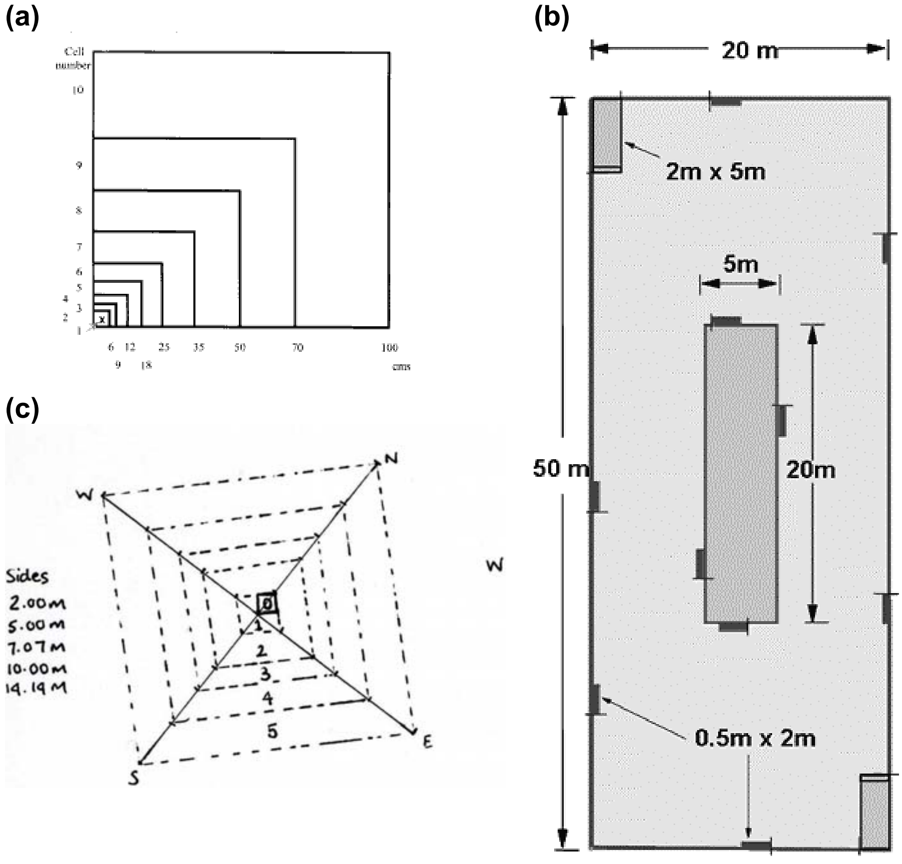


Figure 5.1. Three examples of nested quadrat designs: (a) ADAS (1 × 1m) nested quadrat (Critchley & Poulton (1998), (b) modified-Whittaker Plot (Stohlgren *et al* 1995) and (c) 'Wally plot' used for monitoring vegetation change in Countryside Survey plots (Maskell *et al* 2007). Although designs differ they all use the same approach with recording starting in the smallest cell. Species are only recorded in the cell in which they are first encountered. Therefore, searching in larger cells is confined to the discovery of new species, ignoring those that have already been recorded in smaller 'nests'.

As can be seen from the examples given above, most approaches have been applied at relatively small spatial scales, and so field testing will be required so see to what extent similar methods can be applied at the much larger monad scale. Obvious disadvantages of such an approach might be the extent to which the order of species accumulation is related to the frequency with which a species is encountered at the landscape scale. For example, frequency measures might be less reliable for patchily distributed species, such as aquatics, that would be poorly sampled by this approach (i.e. the likelihood of occurring in the first nest would be much lower than for species occurring in much more extensive habitats such as grassland, woods, etc.). Also the method might be seen as ‘too complicated’ by volunteers not familiar with more systematic sampling techniques.

5.3.4 Transects and random walks

Transects provide an efficient surveillance technique for monitoring changes in the abundance of mobile organisms that are sparsely distributed (and range widely) at a landscape scale. However, they are less effective for monitoring changes in the abundance of plants, because of the ubiquity of plants in most landscapes, the vastly greater numbers of species involved, and their relative immobility which means that changes occur over much longer time periods. Having said that, it should be acknowledged that virtually all *ad hoc* plant recording is transect-based, whereby volunteers attempt to sample as broad a range of habitats present within the recording unit regardless of how large or small.

Recording transects through and around a defined area (sometimes termed ‘random walks’) are an extremely attractive way to survey, especially in England and Wales where land access issues can be a considerable barrier to participation. Such walks are enjoyable, allow a wide variety of habitats to be encountered, and can form part of both everyday activities (e.g. dog walking) or occasional events, perhaps structured around group and family activities. This approach can considerably increase not only the overall participation in surveys of this type, but greatly improve the likelihood for annual repeat survey of individual squares (the walk becoming an annual event).

Recording along transects can either be continual, within sections, in regular sub-plots nested within sections or different habitat types along the route. Walks can be structured as entirely random or using more structured routes, such as the closest approximation to a 'V' or 'W' through the square. Alternatively, transects could be confined to specific habitat types, possibly using fixed ‘V’ or ‘W’ walks similar to those currently used for monitoring the condition of vegetation on conservation sites. The location and width of the route should be preferably fixed although this is possibly the weakest aspect of the method. More value would be attained from the survey if the same walk was repeated each year, so a careful record of the route of the walk would also be required.

The main advantage of transects are that the routes are fixed and therefore that the area sampled is comparable between surveys. The major disadvantage is the general confinement of routes to public rights of way, at least in non-open access areas of England and Wales unless access permission is sought (open access throughout Scotland allows such walks to be undertaken easily in most non-urban areas). Clearly this could introduce a bias in the recording towards linear features such as paths, field margins, etc. Another drawback is the difficulty in relocating and defining the width of the transect, particularly in unenclosed habitats such as moorland.

It would be extremely valuable to field test transect methods, comparing results from squares accessed by public routes and the same squares where private access has been sought.

5.4 Survey methods

In this section we present possible survey methods to record changes in frequency within a national sample of 1km squares. The methods proposed are intended to be exactly the same regardless of whether recording all species in a monad or just a targeted subset. The recording frequency will clearly depend on the level of recording undertaken (Level 3 or 2/1). If recording a subset of target species Methods 2-6 could be run annually. If recording all species all methods could be run every 5 years but with Level 1/2 species being recorded in the intervening years.

To illustrate each method we have used an example monad **TL1389** (Figure 5.2) to simulate how each approach might be implemented in practice. This is a ‘typical’ lowland monad to the east of Peterborough, Cambridgeshire. The land cover is predominantly arable, but includes large blocks of woodland (one of which is ancient), as well as semi-improved and improved neutral grassland. Smaller habitat features with a distinctive flora include road verges (recently invaded by halophytes), boccage around buildings and farms, marshes, ditches/streams, hedgerows and two ponds. Overall the square has an unremarkable flora with few rare species. In 2003/04 a rapid assessment of the square as part of the BSBI Local Change survey recorded 137 taxa although undoubtedly more are present. Of these it is estimated that circa 90% occur in less than a quarter of the 100 hectare grid cells that the square comprises. The sample design for each method is given in Figure 5.3.

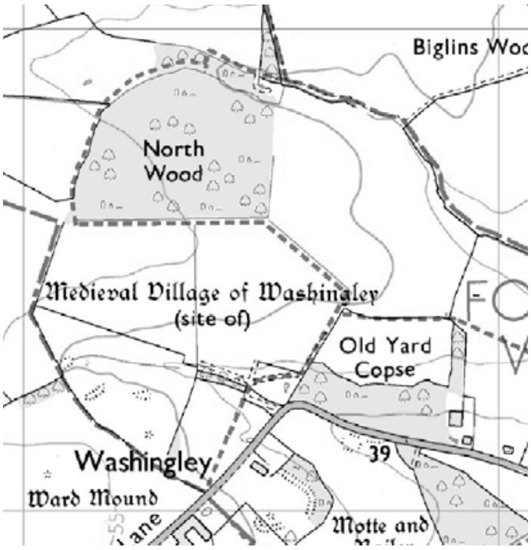


Figure 5.2. Example monad (TL1389) used to illustrate the methods presented in this section.

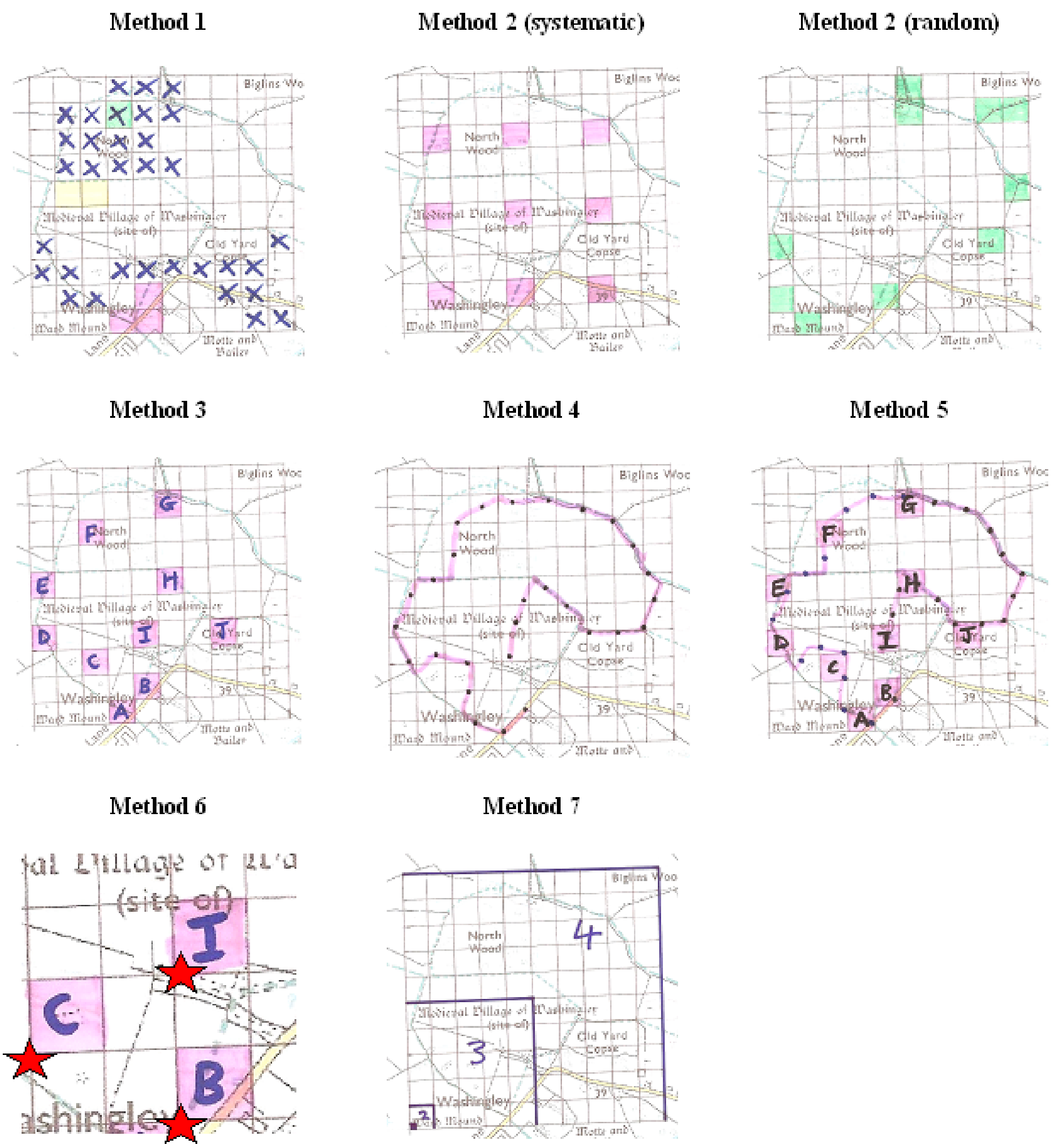


Figure 5.3. Possible approaches to monitoring changes in the frequency of plant species within a national sample of 1km squares. Under Method 1 the colours denote the following species: green – *Orchis mascula*, yellow – *Kickxia elatine*, pink – *Puccinellia distans*, blue – *Fraxinus excelsior*. Under Method 3, 5 & 6 the different letters denote different habitats.

5.4.1 Method 1: Whole square survey

Using this method the square is sampled by visiting a sample of all major habitats present, as it would not be feasible to visit every 100m grid cell in the time available. This route would then form the baseline and would be fixed for all future visits (although there would be no restriction of the width or exact position). For all species or targets encountered frequency would then be estimated using the following frequency classes once the square has been completed: 1, 2-5, 6-25, 26-50, 51-100% of grid cells. Up to four main habitat(s) are also recorded from a predefined list. Recorders note the length of time spent recording and the route taken. The estimated time needed to complete this square would vary depending on experience but is estimated at around 5 hours for an experienced botanist.

The main advantage of this method is its simplicity and comprehensiveness, as an assessment is made for all habitats and species. This approach, is also probably the most effective method for detecting change for rarer species, such as *Orchis mascula* (green squares), *Kickxia elatine* (yellow squares), *Puccinellia distans* (pink squares), the distributions of which are shown in Figure 5.3.

However, this method has many disadvantages. First, results will obviously be biased by the route the surveyor takes which will be difficult to retrace even if maps of the route taken are available. Second, the difficulty is in estimating overall frequency given that many areas remain 'unsampled'. In combination these might mean that 'changes' are artefacts of recording. Third, it would be difficult to measure changes for widespread species because frequency categories are so large (e.g. *Fraxinus excelsior* shown on map). Finally, it would be difficult to attribute changes to habitats or drivers due to a lack of information on species locations.

5.4.2 Method 2: Random/systematic large plots

Record the presence of all species or targets in 10 systematically or randomly located 100m grid cells with an option to carry on recording presence of all other species or targets in the rest of the square. Random grid cells would be chosen and mapped prior to survey to save recorders time in locating squares. For all species or targets recorded, note up to four main broad habitats in each cell. This method would take approximately 6.5 hours to complete (10 × 30 minutes per cell + 1.5 hour to locate cells) with an additional two hours to survey the rest of the square if required. Table 5.2 gives the number of plots sampled per habitat using these two approaches.

Table 5.2. The number of systematic plots recorded in TL1389 using Method 2 and an illustrative random draw

Broad habitat	Systematic	Random	Broad habitat	Systematic	Random
Arable	5	7	Ditch/stream	1	2
Woodland	3	2	Pond	1	2
Hedgerow	3	5	Neutral grassland	1	1
Road verge	2	1	Boccage	1	1
Improved grassland	1	2			

This method provides the most objective assessment of change because it samples species/habitats in an unbiased manner. Plots would also be fixed and therefore easy to relocate. The main disadvantages are that uncommon habitats and species are likely to be missed. Also some plots are likely to be inaccessible (e.g. centres of arable crops). However, possibly the most important disadvantage is that a high proportion of cells are likely to be within apparently 'boring' habitats such as arable land, improved grassland and degraded moorland,

and therefore unattractive to participation by volunteers (but crucial for reporting the effects of policy initiatives such as agri-environment schemes).

5.4.3 Method 3: Large habitat plots

This method using a similar approach to Method 2 but attempts to record the presence of all species or targets in ten 100m grid cells stratified by habitat. These are selected by the surveyor as representative examples of Broad Habitats, from a predefined list, from within the square. Where habitats are smaller than a hectare (e.g. streams, hedgerows, flushes, etc.) only species in the target habitat are recorded as well as an estimate % of cell occupied (i.e. sampled). This method would take approximately the same time to complete as Method 2 (6.5 hours). Whether the frequency of every species within each cell could be estimated using the scale given under Method 1 in the time available requires further field testing.

As with Method 2 this provides an objective assessment of change, but differs in sampling a greater range of habitats that would be more appealing for volunteers to survey (i.e. they would be more likely to see target species as well as others that are uncommon). However, it suffers from a number of weaknesses. First, the sample would not be representative of the extent of habitats in the square as rare habitats would be sampled in the same way as common ones. Further bias would inevitably occur as a result of allowing surveyors to select examples as ‘more interesting’ examples are likely to be selected. A further bias would be introduced by including small habitat patches, although field testing would be required to assess how these affect overall results.

5.4.4 Method 4: Large transects

All species or targets are recorded along a continuous 2km, fixed-width (2m) transect route (or 2 × 1km). This route is selected by the surveyor and should aim to sample a range of habitats present in the square. For the purposes of recording frequency, the transect is divided into fixed sections of 100m (though this needs field testing). The occurrence of all species and targets along with up to four habitat types are recorded for each section. This would take approximately 5 hours to record (20 sections × 15 minutes) though this would require field testing.

Like the previous methods a fixed sample would allow an objective assessment of change that would be appealing to volunteers. Furthermore, this approach could be combined with transects recorded for other taxonomic groups, allowing multi-taxon comparisons to be undertaken. The main disadvantages of this method would be the inevitable bias of transect routes towards linear features and the difficulties in retracing routes through unenclosed habitats. As in Methods 2 and 3, large areas of the square remain ‘unsampled’. A further problem would be the temptation for recorders to include species from outside the transect. Although these data could be recorded it is difficult to know how these would be analysed.

5.4.5 Method 5: habitat transects with small plots

Method 5 has been split into two methods, 5a and 5b, following discussions at the workshop.

Method 5a is a modification of Method 3; rather than recording the presence of species across all habitat patches within the 100m grid square, a single fixed transect is placed within a representative area (maximum of 10 per square). The actual number, size and orientation of these transects requires field testing but could include straight, ‘V’ or ‘W’ shaped 100 × 2m transects depending on the dimensions of the habitat patch.

During the workshop, the map for Method 5a was interpreted as a "walk" through the grid square, recording as you go (possibly in sections) but stopping to record small quadrats in each habitat in the square (maximum of 10 per square). This method, which is a combination of Methods 3, 4 and 6, was popular and considered to be attractive to volunteers. As with method 3 grid cells are stratified by habitat as representative of examples of Broad Habitats within the monad, these grid cells are linked by a transect route, or random walk as in method 4. The surveyor would record along the transect/walk and within grid cells using small plots (quadrats) as outlined in method 6 below. The advantage of method 5b would be combining an enjoyable activity (walk) with detailed recording (plots), however disadvantages include the time such a survey would take, depending on detail of walk/transect recording, and disadvantages associated with quadrat recording which are outlined in more detail below.

5.4.6 Method 6: Small plots (quadrats)

Quadrats could be recorded as an additional module used on Methods 2 to 5 (i.e. within grid cells, transects, transect sections) in order to record the abundance of a subset of species recorded in the larger plots. These could be located as follows:

- **Method 2** – the SW corner of random/systematically located 100m grid cell;
- **Method 3** – the centre of a representative patch of habitat for which the 100m cell was selected;
- **Method 4** – the centre/corner of the transect at the start of each section;
- **Method 5a** – the centre/corner of the start of the transect within the habitat patch;
- **Method 5b** – the centre of a representative patch of habitat for which the 100m cell was selected along the long transect/walk.

The actual size and dimensions of the quadrats would be habitat specific but we would envisage 1-2 square metre plots in most habitats. Plots would be fixed and relocated using a handheld GPS and sketch maps. Abundance would be estimated on the DOMIN scale. The time needed to complete 10 plots would be around 5 hours and so combining this approach with Methods 2-5 would probably not be achievable in a day.

The main advantage of this approach would be the very clear signal of change from year to year. However, this is likely to be an unattractive method for many volunteers. Other disadvantages include the large sample size needed to ensure sufficient replication; inaccessibility and difficulty of relocating quadrats in some habitats (e.g. in crops, ravines, etc.); and that the sampling approach is not achievable in a day when combined with Methods 2-5. Overall this approach might add little for a high cost.

5.4.7 Method 7: Nested plots

In this method the recorder records the first occurrence of all species or targets within nests of increasing size. Although this requires field testing, a simple version would include nests of 10m, 100m and 500m within a single kilometre square starting in the southwest corner (Figure 5.3). The habitat of each species would also be recorded in the first nest in which it occurred. The position of the nests need not be fixed and could be rotated around the square corners, although the implication of this would require field testing. Because of its novel nature it is difficult to estimate the time needed to complete a nested survey but we have no reason to believe it would be longer than the 5-6 hours needed to complete Methods 2-5.

The main advantage of this method is the relative simplicity by which detailed information on frequency are derived. It also provides a complete census of the whole square using an efficient search method whereby a recorder focuses on the discovery of new species rather than re-recording the same species in multiple samples. One disadvantage is the placement of the smallest nest which might be difficult to relocate without permanent markers, or inaccessible. The positioning of the smallest nest would therefore require field testing. Other potential difficulties include navigating the whole square to ensure all new species are picked up (for levels 1-3) and the time this would take to complete, and as such its attractiveness to a range of volunteers.

5.5 Derived statistics and policy relevance of each method

It must be understood that it is very difficult to link specific environmental drivers to each potential method. The degree to which environmental changes can be detected through a plant surveillance scheme are intimately linked to the quantity and diversity of data which can be gathered, which in turn requires field testing of potential methods and further statistical analysis.

Additionally, species indexes referred to below can be interpreted against a range of environmental drivers (relevant to Objective 2 of the TBSS) based on linking these with species and traits, and coupled with additional parameters recorded in transects or plots.

Method 1 – whole square frequency estimate (out of a possible 100) used to derive a national index for a species based on the monads sampled for a given species in a given year. Method would be good at reporting trends for a large suite of species including those that are more restricted nationally (assuming all species are recorded at Level 3). This method meets the requirements of objective 1 of the TBSS and potentially objective 3.

Method 2 – whole square frequency estimate (out of possible 10) used to derive a national index for a species based on the monads sampled for a given species in a given year. Possibility of calculating indices by habitat although complicated by the fact that more than one can be recorded per square and therefore that any observed decline/increase cannot be attributed directly to a single habitat type (Method 3 overcomes this problem by nesting plots within habitats). Both approaches (systematic and random) would be good at reporting trends for common species occurring in widespread habitats, but only limited coverage would be achieved for more restricted species and habitats (even if encountered as sample sizes are likely to be small). The policy relevance would therefore be in reporting trends in the wider countryside. This method meets the requirements of objective 1 of the TBSS.

Method 3 - whole square frequency estimate (out of possible 10) used to derive a national index for a species based on the monads sampled for a given species in a given year. This method has the advantage that indices can also be derived for different habitats. The number of samples within each habitat type however would be critical and therefore an assessment is needed to simulate the likely replication that this approach might achieve given different sample numbers. By attempting to sample a range of habitats this approach would inevitably bias the survey towards rarer (semi-natural) habitat types, thereby improving the chance of encountering more restricted species. The policy relevance would therefore be in assessing the quality of habitats under policy initiatives (such as AE Schemes, protected areas, etc.) versus the wider countryside as well as (potentially) the changing status of priority species (e.g. BAP, Red List, etc). This method meets the requirements of objective 1 of the TBSS and potentially objective 3.

Method 4 - whole square frequency estimate (out of possible 20 transect sections) used to derive a national index for a species based on the monads sampled for a given species in a given year. As in Method 2 the possibility of calculating indices for habitats is complicated by the fact that more than one can be recorded per transect section and therefore that any observed decline/increase cannot be attributed directly to a single habitat type. Also as with Method 2 this approach would provide trend information for common species occurring in widespread habitats, but only limited coverage would be achieved for more restricted species and habitats (even if encountered as sample sizes are likely to be small). In addition there is likely to be a bias to habitats associated with linear features (i.e. that allow access), unless access permission agreed, and an under-representation of species of unenclosed habitats that are less accessible (e.g. moorland, water-bodies, etc.). The policy relevance would therefore be in reporting trends in the wider countryside. This method meets the requirements of objective 1 of the TBSS.

Method 5a – very similar to Method 3 but would provide indices for fewer species as only those in a sub-sample of the entire area of habitat are being recorded. A similar method is currently used in England as part of the assessment to establish the ‘condition’ of designated sites. Potentially highly policy relevant for measuring habitat condition but probably not so effective for establishing trends for individual species due to the limited area sampled (and therefore small sample size for most species other than those that are common and widespread). This method meets the requirements of objective 1 of the TBSS.

Method 5b – very similar to Method 4 (assuming frequency is recorded in 20 transect sections) but additional small habitat plots overcomes the problem of not knowing in which habitat changes are occurring (assuming a species for which a trend has been detected at the transect level is also present in one of the quadrats). Like Method 5a potentially highly policy relevant for measuring habitat condition but probably not so effective for establishing trends for individual species due to the limited area sampled (and therefore small sample size for most species other than those that are common and widespread). This method meets the requirements of objective 1 of the TBSS.

Method 6 – provides abundance measures that can be used to derive a national index for a species based on the monads sampled for a given species in a given year. Only likely to cover a subset of the species for which frequency measures are available at the larger plot, habitat patch or transect scale and tending to be the commonest species in the sample as a whole (due to small area sampled). Policy relevance therefore probably limited to reporting trends in the wider countryside when combined with Methods 2 and 4 but probably more focussed on more restricted species when combined with Methods 3 and 5 (i.e. when stratified by habitat type). Overall this method might add little for a high cost.

Method 7 – the nest positions (i.e. first, second, etc.) are used to derive a national index of frequency using the assumption that a species present in Nest 1 is more abundant than a species first recorded in Nest 3, 4, etc. Clearly this will not always be the case and field testing will be required to assess the relationship between position in nests and abundance at the monad scale. As with Methods 1, 2 and 4 limited assessments could be carried out by habitat. Potentially a promising approach that overcomes many of the weaknesses of some of the other methods such as bias to semi-natural habitats, and the unappealing nature of recording common species in small plots in 'uninteresting' habitats. Overall the approach is highly policy relevant because of the range of the scale of the sampling allows reporting of trends for common, widespread species as well as more restricted taxa present in the rest of the square. In addition responses to policy initiatives could be measured by comparing trends in the smaller nests only, where a subset lie within nature reserves, AE scheme agreement holdings, etc. although comparisons would be based on changes in frequency across the sampling domain rather than actual changes in abundance (e.g. under Method 6). This method meets the requirements of objective 1 of the TBSS and potentially objective 3.

5.6 Conclusions

The methods presented in this chapter were assessed by specialists at a workshop where participants were asked to select their two preferred approaches. Methods 3, 5b and 7 were the most popular.

We recommend that aspects of all three approaches are investigated further through field trials and statistical analyses and simulations. It is recommended that the different methods are trialled in the same squares and that methods are compared when restricted to public rights of way and when private access has been arranged. The two key questions to address during these trials are:

- What is believed to be achievable, and what could it deliver?
- What does increasing the number of different habitats sampled provide in terms of analysis (assuming that we do not need to estimate habitat area but we do need to estimate changes in habitat quality).

6 Defining optimal survey times for UK Broad Habitats

6.1 Habitat phenology

In order to define the optimal times to survey species in British habitats we analysed the published flowering data for all native and archaeophyte taxa. These data were initially used to derive the 'phenology' of each of the 20 UK Broad Habitats (BHs) given in PLANTATT (Hill *et al* 2004). BH23 – Inshore sub-littoral sediment – was excluded from this analysis as it only includes one species - *Zostera marina*. In carrying out these analyses we made the following assumptions that are likely to be false for a number of species (though justifiable in this context):

- All species produce flowers in the UK and this period is distinguished from fruiting although clearly the two overlap. In some databases 'flowering' has been used in the broad sense to include fruiting where the two are difficult to separate (e.g. oraches, docks, grasses and sedges) although for sedges, Jermy *et al* (2007) recently attempted to differentiate the two. Ferns and fern allies pose additional problems and so we have used Page (1997) who gives the months when spores are produced;
- To our knowledge, no British species fail to produce flowers, although a small number are cleistogamous (e.g. *Epipactis, Viola*) or rarely produce flowers in the UK (e.g. *Pyrola media, Carex vaginata*), which would obviously affect detection rates during any surveillance scheme;
- Species are easiest to identify when in flower and/or fruit. This is probably not true for most pteridophytes although reproductive structures are important identification and detection characters for many taxa;
- Where a month is given, the species is assumed to be in flower for the entire month. Most databases take this approach which is probably acceptable given geographic variation in earliest and latest flowering months across the country;
- There is little geographic/habitat variation in the timing of flowering across GB. Clearly this is not true although as stated above, a broad definition is probably the best we can currently hope for.

6.1.1 Method

For each species, flowering months were extracted from the ECOFLORA and Comparative Plant Ecology traits databases (Fitter & Peat 1994; Grime *et al* 2007) for all native and archaeophyte species listed in PLANTATT (Hill *et al* 2004). Where values were missing or considered unreliable, values were extracted from recent botanical works. These included Cope and Gray (2009) for grasses, Page (1997) for clubmosses, ferns and horsetails and Streeter *et al* (2009) for all other species.

Flowering months were available for 1380 native (including 41 'native or alien') and 150 archaeophyte taxa. The total number of species flowering in each month was calculated for each Broad Habitat, with each species potentially being assigned to up to four Broad Habitats, as given in PLANTATT. The numbers of species flowering in each month was then converted to the proportion of all species present in the Broad Habitat, and then plotted against each month.

Because uncommon species are unlikely to be recorded in a national sample survey we carried out the same analyses after removing species present in fewer than 750 hectads in GB. This reduced the total to 685 taxa.

6.1.2 Results

Mean flowering dates for all Broad Habitats are shown in Figure 6.1. With the exception of Broad-leaved Woodland and Saltmarsh, which had the earliest (early June) and latest (late July) mean dates respectively, the mean dates were restricted to mid-June to mid-July. With the exception of Saltmarsh, the exclusion of taxa in <750 taxa had very little effect on mean dates for most Broad Habitats.

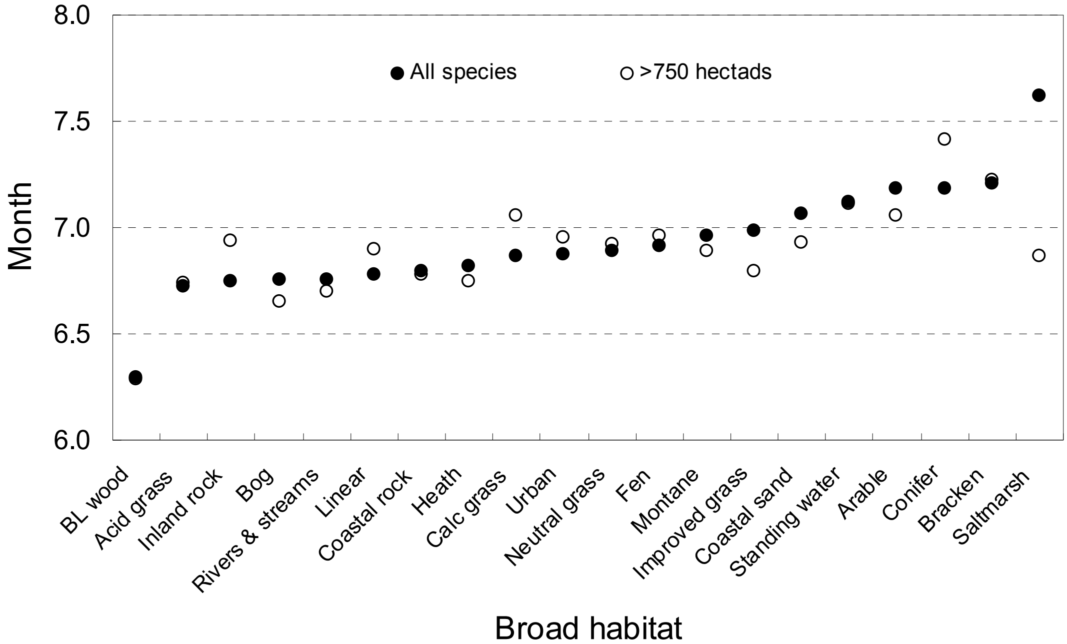


Figure 6.1. Mean flowering dates for British Broad Habitats based on the flowering times of native and archaeophyte taxa. Means are shown for all species (black circles) and species recorded in more than 750 hectads (open circles). Error bars excluded for clarity.

Figure 6.2 shows the flowering profiles for all Broad Habitats based on the flowering dates for all species (blue), and only those recorded in more than 250 hectads (red) and 750 hectads (green). Clearly most Broad Habitats display a pronounced ‘peak’ in flowering during the mid-summer though Broad-leaved Woodland and Acid and Neutral Grassland are skewed towards the spring and Heath, Bracken and Coastal skewed towards the late summer. In comparison, a few Broad Habitats display a distinct ‘plateau’ due to the long period over which species flower (e.g. Broad-leaved Woodland, Bog).

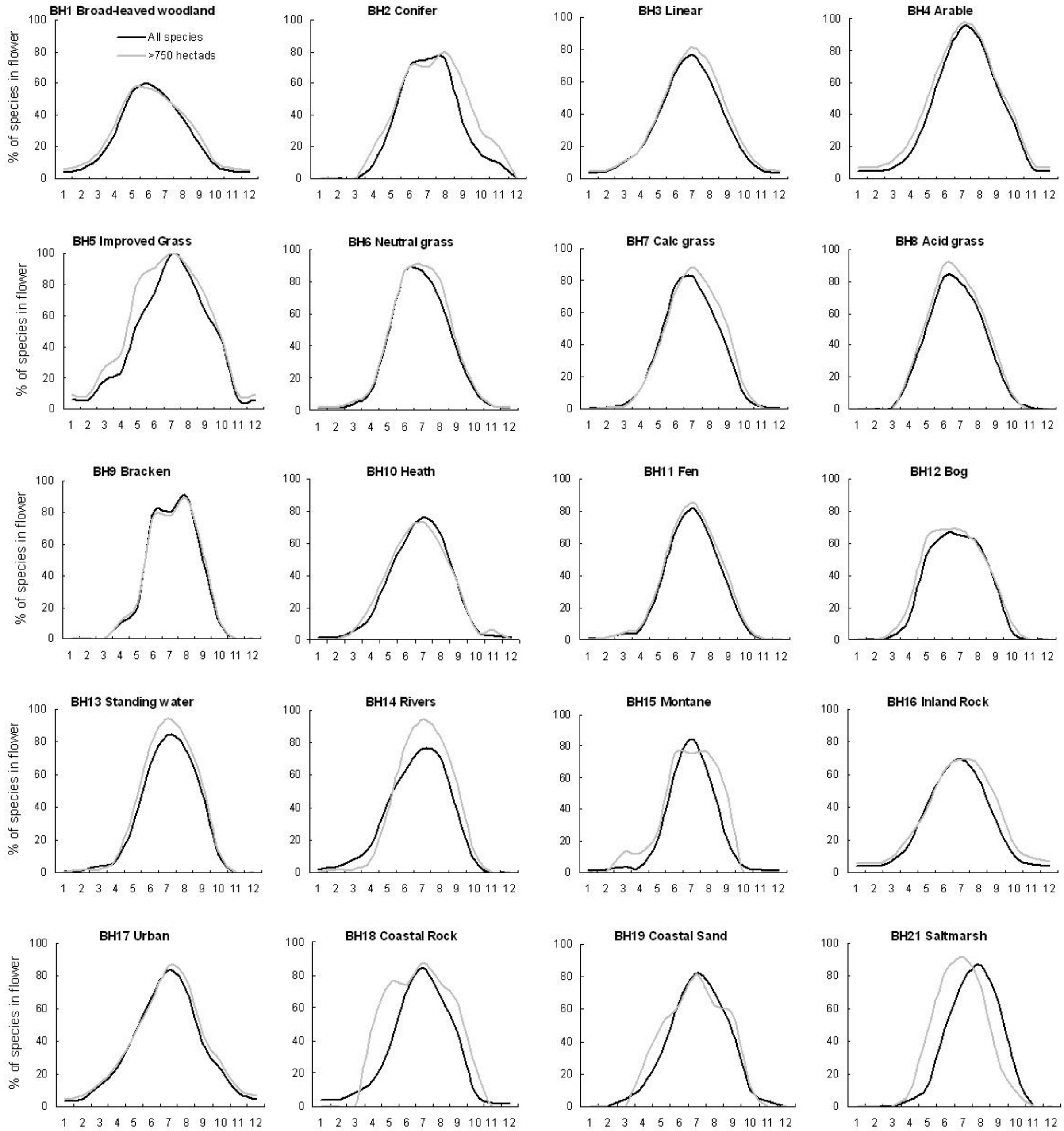


Figure 6.2. Phenology of UK Broad Habitats based on the monthly flowering times of native and archaeophyte taxa. Graphs are plotted for all species (black line) and species recorded in greater than 750 hectads (grey lines).

It is also noticeable how 'low' some of the peaks are (e.g. Broad-leaved Woodland, Heath, Bog, Inland Rock). For these habitats surveys during peak flowering could potentially miss between 20-40% of the species likely to be present. Therefore, for these habitats there is a strong argument for multiple visits or at least careful checking of flowering times during target species selection. In contrast a single (peak) visit is likely to detect >80% of species for most other habitats and >90% for Arable, some grasslands and aquatics habitats (for species in >750 hectads).

The exclusion of species in less than 750 hectads had no significant effect on mean flowering date, with the exception of the Saltmarsh Broad Habitat for which the mean flowering date was significantly earlier when species in less than 750 hectads were excluded (two-sampled *t*-test: $T = 3.10$, $P = 0.003$). This is shown more clearly in Figure 6.3; here the differences between the mean flowering for all species and those in greater than 250 and 750 hectads are presented. All other Broad Habitats only saw a shift in the mean flowering date by less than 0.2 of a month. Those with larger mean shifts tended to be Broad Habitats containing fewer species (i.e. Improved Grass, Conifer, Coastal Sand).

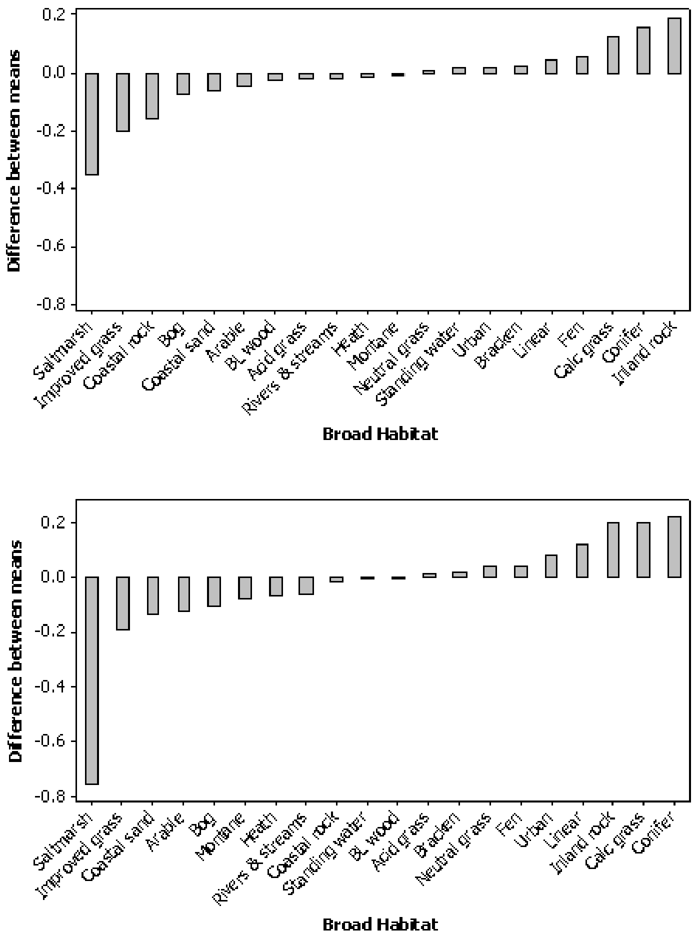


Figure 6.3. The difference in mean flowering dates for UK Broad Habitats. Comparison of means derived for all species and species in more than 250 hectads (top) and 750 hectads (bottom).

6.2 Conclusion

Assuming a single visit is made, the peak flowering month would seem to be the optimum time to survey a habitat as it gives the greatest chance of sampling the greatest range of species. Using this approach, and the data presented in this chapter, we propose the optimum time to record Broad Habitats in Table 6.1. Clearly late June/early July would be the optimal to survey most habitats. An earlier visit would be required for Broad-leaved Woodland, Acid Grassland and Bog whereas a later visit in August would be most likely to record most species in Saltmarsh and Bracken. Multiple visits might be required to ensure sufficient species detection in some habitats (denoted by an asterisk in Table 6.1).

With the exception of Saltmarsh, the proportion of species recorded (all, >250 10km, >750 10km) has no significant effect on these optima, although for Broad Habitats with small numbers of species, there is a small amount of variation around the median month (e.g. Conifers, Bog, Montane).

Table 6.1. The optimum months to survey UK Broad Habitats based on the numbers of higher plants in flower in any one month. Where 2 or more values are given, months had identical proportions of species in flower. The earliest and latest months for survey are also given based on the 25th and 75th percentiles of species in flower (using only species recorded in >750 hectads). These provide the range within which any additional surveys should take place.

Broad habitat	Optimum month to survey			Survey limits	
	All species	>250 10km	>750 10km	Early	Late
Broad-leaved woodland*	6	6	6	5	8
Conifer woodland	7 or 8	6, 7 or 8	8	6	9
Linear habitats	7	7	7	6	8
Arable	7	7	7	6	8
Improved grassland	7	7	7	5	8
Neutral grassland	6 or 7	7	7	6	8
Calc grassland	7	7	7	6	8
Acid grassland	6	6	6	6	8
Bracken	8	8	8	6	8
Heath*	7	7	7	6	8
Fen, marsh and swamp	7	7	7	6	8
Bog*	6 or 7	6	6 or 7	5	8
Standing water	7	7	7	6	8
Rivers & streams	7	7	7	6	8
Montane	7	7	6 or 7	6	8
Inland rock*	7	7	7	6	8
Urban	7	7	7	6	8
Coastal rock	7	7	7	5	8
Coastal sand	7	7	7	6	8
Saltmarsh	8	7	7	6	8

* Habitats possibly needing multiple visits to ensure detection rates are >80%.

Table 6.1 also provides an indication of the range of months in which earlier or later visits should be undertaken. These are the months in which the 25th and 75th percentiles falls based on only species recorded in greater than 750 hectads. An earlier visit in May or June is very much dependent on the habitat whereas in all cases a later visit would be carried out in August.

Multiple visits are only likely to have a detrimental effect on vegetation structure, and possibly species composition, on small plots (<10m), or in a wetland habitats that are very sensitive to repeated disturbance (e.g. flushes, mires, bogs, fens, marshes). Such considerations would need to be taken into account during the design of any scheme and/or tested through further field survey.

Although these analyses suggest the optimum month(s) to survey a habitat, in reality other factors may be more important in determining the actual timing of surveys. First, surveyors are unlikely to know, in the first year at least, which habitats are present in their sample squares and therefore plan their visits accordingly. Second, many botanists will be able to identify species not in flower, especially later in the season when fruits are present and therefore feel less constrained to the months indicated. Finally, it may not be possible to survey certain squares at the appropriate time for purely practical reasons. In conclusion, with the exception of Broad-leaved Woodland and Saltmarsh, a late June/early July visit would seem optimal for the majority of British Broad Habitats preferably with an early and/or later visit in the months immediately preceding and following this optimum.

7 Recording additional parameters

7.1 Introduction

Additional attributes might be recorded where they provide local level explanatory information which could not be derived from other national datasets (e.g. Land Cover, Agcensus, canopy structure from LIDAR imagery, nitrogen deposition, etc.). These should only be selected where they relate to explanatory factors (e.g. eutrophication, climate change, etc.) that might otherwise confound analyses of multiple drivers (a good example of a very simple measure would be average sward height in grasslands). Such attributes might also provide extra analytical power for secondary analyses such as predictive modelling of the likely impacts of new crops. Such information might also allow an assessment of habitat condition to be made and potentially provide valuable information useful for other taxonomic groups. Such attributes are not easy to develop however, and much depends on the exact survey method employed. Indeed many attributes considered in this section can only be usefully assessed in very small sample plot areas (e.g. plots <10m), and therefore might not be relevant to a scheme employing larger plot sizes. Crucially, additional attributes can considerably increase the survey time required to complete sampling and therefore might not be practical within a voluntary scheme. These can quickly turn an attractive, rapid survey into a long and laborious one.

Potential additional attributes are presented in this section regardless of their appropriateness for a new surveillance scheme; their implementation in the final scheme will be dependent on the practicalities of collection and successful field-testing.

7.1.1 Vegetation measures

i Presence/absence

The presence of a species is the simplest form of vegetation assessment, and is relatively easily undertaken by surveyors of all abilities. The main benefit is the speed at which species are recorded, thereby allowing recorders to move on to (i.e. look for) the next species as soon as the targets has been located. Their main disadvantage relates to plot size. For practical reasons larger plots are more difficult to survey due to terrain, time available, access, habitat diversity, etc. Conversely, in smaller plots, recording can be more difficult if surveyors attempt to identify plants that are not in flower.

ii Cover-abundance

The **percentage cover** (vertical projection) of a species is commonly used to estimate abundance within small plots. However, estimates vary greatly between surveyors especially where the vegetation is multi-layered. Other difficulties include the identification of vegetative material (which can sometimes be the dominant species) and whether to include species ‘rooted’ outside the quadrat. Cover-abundance categories reduce some of the variation between surveyors and speeds up the process of estimation considerably. The most widely used is the Braun-Blanquet or Domin scale (<1%, 1-4%, 5-9%, 10-24%, 25-32%, 33-49%, 50-74%, 75-94%, 95-100%) which is thought to take about a fifth of the time it takes to record ‘exact’ cover. If required these categories can be transformed back into percentages using the ‘Currall 2.6’ transformation (Currall 1987).

iii Frequency

Another way to estimate abundance is to measure frequency (presence) in cells. This is the system currently used by ECN where it is thought to be much more reliable than % cover and Domin for monitoring vegetation change. However, it requires a large number of cells to be surveyed thereby significantly increasing the time taken to complete a sample. Also surveyors differ in only counting species that are rooted in rather than ‘covering’ cells (*Calluna vulgaris* in moorland is a good example of where this method provides only a very poor assessment of overall cover-abundance).

iv DAFOR

The DAFOR scale (Dominant, Abundant, Frequent, Occasional, Rare) has been used to measure the abundance of species at a variety of scales although it was originally designed for recording abundance in small plots. Although largely subjective, it provides broad categories that are relatively easy and rapid to interpret. Visual representation could be used to help those not familiar with the scale to use it in the field.

7.1.2 Structural measures

Various structural measures that help to indicate habitat quality are routinely used in rare plant population surveys (e.g. Plantlife’s Back from the Brink, the BSBI’s Threatened Plants Project, Common Standards Monitoring). Such measures can provide useful information on the status of the habitat, including levels of management and disturbance.

i Vegetation height

Research has shown that sward height calculations can vary markedly depending on the method used (e.g. direct measurement, sward stick, drop-disc, etc.; Stewart, Bourn & Thomas 2001). However, it is unlikely that direct measurements would be practical in a large participation scheme given the cost and practicalities of supply and implementation. As with cover-abundance an alternative is to use a simplified classification that allows rapid assessment. For example, the BSBI’s Threatened Plants Project uses <10cm, 11-30cm, 31-100cm and > 100cm to indicate the average sward height associated with a range of threatened (target) species. However, complications occur where the vegetation has a mixed structure (e.g. heathland/acid grassland mosaics) or is multi-layered, as in woodland. Another alternative is to carry out a post hoc analysis using the maximum summer (canopy) heights for British and Irish species given by Hill *et al* (2004). However, further research is needed to test how well this method ‘predicts’ the correct vegetation height using real data.

ii Bare ground

The same problems associated with measuring cover-abundance apply to bare ground. While extremely useful, surveyors would have to record either broad categories of bare ground (as above) or using a DAFOR scale in small plots for the information to be of use. Having said that volunteers tend to find bare ground easier to estimate than plant cover, as it tends to form more discrete patches (although complications arise where it is intermixed with litter, mosses, bare rock, gravel, etc.). However, given time constraints it is unlikely to be a priority variable to record in any future scheme.

iii Woody cover

The same problems recording bare ground apply to measures of percentage woody cover. While useful, surveyors would have to record either broad categories (as above) or as a DAFOR scale in a very small plot. Woody cover can be difficult for volunteers to interpret (e.g. for instance would all scrambling shrubs, such as brambles be included) and clear guidelines on what and how to record would be needed.

iv Grass:herb ratio

This can be very a very useful measure and volunteers find it relatively easy to estimate in smaller plot sizes, especially if stratified into broad categories (e.g. 0:100, 25:75, 50:50, 75:25, 100:0). The main problem is heterogeneity across stands of vegetation and the need for a larger number of samples to derive an average ratio for a particular habitat.

7.1.3 Habitat management parameters

i Grazing and other management

Information on grazing and other management is often vital for explaining vegetation change, measuring habitat quality and also assessing the success of policy initiatives, especially if measured over a period of time. For most habitats the presence of grazing animals is probably the most important and is relatively simple to judge (especially in grassland habitats, less so in woodland). However, an indication of grazing intensity is much more informative. The BSBI's Threatened Plants Project defines broad categories (high, moderate, low, none) which are easy to interpret by volunteers and supported by an estimate of sward height. The extent to which the habitat is under- or overgrazed is also captured within a list of potential threats affecting populations of the target species being surveyed (see below). Plantlife's Flora Guardian programme provides volunteers with a check list of land management issues to record on site.

Many other parameters could potentially be recorded to indicate important activities that are likely to explain long-term vegetation change. Most are habitat specific and could use the same scale (i.e. high, moderate, low and none). A few examples include:

- Burning (moorland);
- Cutting/mowing (grassland, road verges, heathland);
- Coppicing/planting/restocking (woodland and hedgerows);
- Ditch-clearing/drainage (wetlands);
- Control of 'wild' animals (e.g. rabbits/deer).

ii Habitat quality

The overall quality could be monitored with a subjective assessment of its condition. Recorders could be asked to assess the quality of each habitat in their square (e.g. excellent, good, reasonable, poor, destroyed etc.). Although very subjective and probably not at all suitable for volunteers, this could prove useful if alternative (and more direct) measures are too time-consuming and difficult to apply in larger sample areas.

iii Threats

An assessment of threats (noted or perceived) could provide valuable information to support analyses to determine the key drivers of change in different habitats. The following list is used in the BSBI's Threatened Plants Project and could form the basis of any future assessment:

- Afforestation including 'farm woodland' tree planting;
- Agricultural improvement (e.g. ploughing, drainage, reseeding, etc.);
- Burning (e.g. accidental, 'controlled' on moorland, etc.);
- Lack of management (e.g. general dereliction, cessation of former traditional use);
- Invasive alien or native species (these could be included in the target species lists);
- Mineral extraction (e.g. hard quarrying, gravel extraction, etc.);
- Pollution/eutrophication (e.g. slurry or manure use, fertiliser drift, algal blooms in water bodies);
- Herbicide use;
- Recreational activity (e.g. trampling, building of infrastructure, etc.);
- Urban/road development (e.g. road widening/building, housing/industrial developments, etc.);
- Overgrazing or disturbance caused by deer, rabbits (non-livestock).

7.1.4 Attributes valuable for other taxa

Plants provide both shelter and a source of food for the majority of our other terrestrial wildlife. Measures of the availability (presence or abundance) of nectar, seed and other target food plants could prove very useful in linking a plant surveillance scheme with data collected for other taxon groups, especially birds and butterflies.

i Pollen and nectar sources

The availability of pollen and nectar, as a source of food for invertebrates, can be assessed if smaller quadrats are employed. This could be measured by the presence or absence of open flowers, the presence or absence of flowers on particular indicator species (such as certain taxa of the Asteraceae and Lamiaceae), the abundance of all flowers or particular species (either % cover or presence/absence in the four quarters of a quadrat), or the ratio of flowering and non-flowering plants. This will obviously be highly variable and dependent on time of year etc, but could indicate the effects of various management practices in certain habitats (e.g. overgrazing of grassland, mowing of roadside verges).

ii Seed availability

The availability of seed could be assessed in a similar way to nectar by recording the presence of seed heads. Differentiation could be made between monocots and dicots, the presence of specific important species or groups of species (such as Chenopodiaceae), and the differing availability of nuts, large fruit, berries and seeds. Again, the timing of the surveys will have a significant impact on the data collected and repeat visits might reveal interesting patterns of seed or fruit availability.

iii Food plants

For some non-plant taxa, the presence of particular plant species can be important, especially if they are food plants for priority species. An example would be Devil's-bit Scabious *Succisa pratensis* and Marsh Fritillary butterfly. In such cases, it would be useful to ensure the plant species are included in the list of target species. It might even be possible to encourage participants to record the presence of the non-plant species when the plant is recorded. Advice will be sought from other partners regarding the appropriateness of this approach and compatibility of data.

7.1.5 Identifying key attributes

Field testing will be required to determine the key attributes that could provide robust data under the sampling model adopted. The key consideration will be the size of the plot as this will determine the level of detail that can be recorded and the time it will take to complete the survey. Most quantitative attributes (e.g. vegetation height) become very difficult or impossible to assess when plot dimensions exceed 2 × 2m whereas descriptive attributes (e.g. level of grazing) are independent of scale. This does not always mean that quantitative attributes cannot be recorded in larger plots just that different measures are used (e.g. simplified broad categories). Table 7.2 attempts to summarise a possible set of optimal attributes to record and how these could be recorded in small (2m or less) habitat plots and in larger plots.

Table 7.1. Possible set of optimal attributes to record depending on plot size. It is likely that the methods recommended here will employ plots larger than 2 metres and these measures are therefore preferred.

Attribute	Small plots (<2 × 2m)	Larger plots (up to 100 × 100m)
Target species	abundance categories	presence / absence**
Bare ground*	abundance categories	n/a
Vegetation height*	height categories	minimum / maximum height
Woody cover*	abundance categories	abundance categories
Grass:herb ratio*	ratio categories	n/a
Other management (defined)	intensity categories	intensity categories
Disturbance	intensity categories	presence / absence
Important nectar sources for other taxa	abundance categories	presence / absence
Important seed sources for other taxa	abundance categories	presence / absence
Important food plants for other taxa	abundance categories	n/a

7.2 Conclusions

Which attributes can be collected will depend very much on the method of survey chosen. It is recommended that the collection of the attributes in Table 7.1 are tested in the field with each survey method in order to assess how easily the data can be collected, the amount of time they add to the survey time, and the subsequent value of the data collected. It will also be important to examine the variability in data collected between recorders so this can be minimised. Measures of presence/absence are, however, preferred, as these are more reliable and easier to collect than abundance. The number of plots in a 1km square in which a species occurs can provide measures of relative frequency or relative abundance. Finer scale measures of vegetation change (i.e. using measurement in quadrats under $2m^2$) might be difficult given the need for rapid survey using volunteers.

8 Analytical power and sample size

8.1 Introduction

Monitoring and surveillance schemes for plants in the UK routinely gather data from a widely-distributed sample of 'sites' of standard size (hereafter referred to as 'squares'), at which the species present are recorded. Given the difficulty of forming a comprehensive species list for a large site/square, an appealing alternative is surveying in detail a number of much smaller random subsidiary sites within each of these primary units (this equates to Method 2 in section 5.4.2). An illustration of this approach is given in Figure 8.1.

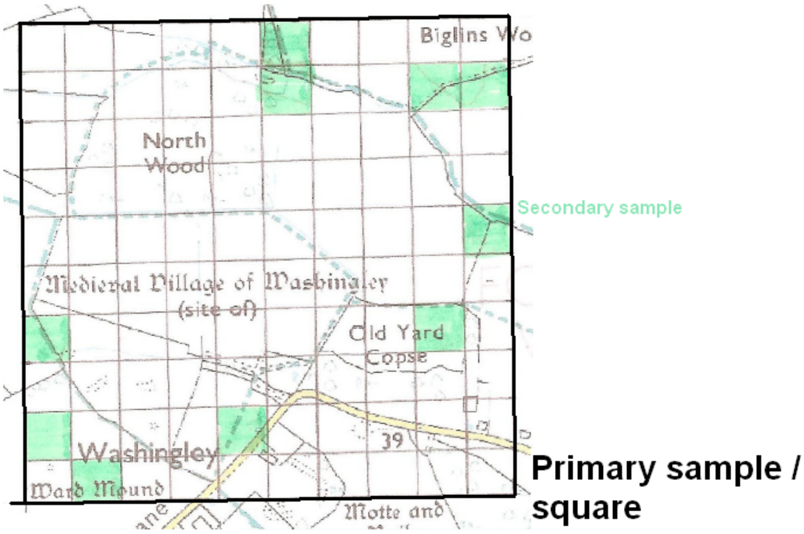


Figure 8.1. Example of a primary site or square (e.g. 1 × 1km) and ten randomly chosen secondary samples within the primary sample.

On a national scale, even relatively widespread species will only be recorded in a small proportion of squares that are sufficiently small for a complete assessment to be made (for example in 2 × 2km squares as described in Braithwaite *et al* 2006). In general a large number of primary squares and an appropriate statistical model are therefore required for accurate and statistically robust identification of species' trends (both increases and declines). The primary purpose of this section is to investigate the size of sample required to detect reasonably substantial and enduring changes in plant abundance over time, via an analysis of statistical power based upon simulated sets of data. As the data are in the form of proportions (% of secondary sampled units occupied) the natural models are binomial in form and easily fitted in widely-available software. Binomial data are also relatively straightforward to simulate, enabling assessment of the power to detect specified changes.

Given the coarse resolution of 'presence/absence' data, and the limitations on the number of secondary samples practical in the available time, such power is likely to be limited for all but substantial changes in relatively widespread species. We conclude this section therefore with additional discussion of possible alternative sampling strategies predominantly arising from discussions at the Salisbury workshop. These strategies require less standard techniques of analysis, and cannot for example be fitted as simple Generalized Linear Models as are those above. Bespoke computer programmes are therefore required to fit the models by maximum likelihood via a numerical maximisation routine. A basic programme of this kind has been written for one of the considered methods, and initial analyses based upon real (Plantlife CPS) data and simulated results are presented here. Although a full-blown power analysis of these latter models is outside the remit of the current project and the time available, discussion is provided regarding the precision of parameter estimates, which is closely linked to the prospects of identifying species trends. These studies adopt the collection of more detailed information which can be expected to translate into greater power and, as such, we believe they are worthy of future research into the viability of their adoption into national botanical surveys.

8.2 Power of models for presence and frequency data

We assume a sampling design based upon randomly selected primary units and, within these, randomly selected secondary units (i.e. Method 2; section 5.4.2). The primary units are assumed too large for complete, accurate monitoring (e.g. $1km^2$), whereas the secondary units are assumed small enough for any species present not to be overlooked (e.g. $1\text{-}10m^2$). The analyses, however, are formally scale-independent and not in any way restricted to areas of these dimensions and therefore the conclusions would hold for the 100 × 100m secondary units of Methods 2, 3 and 5 described in section 5 of this report.

Initially we assume that 100 primary units are sampled, and within each of these five secondary units. Data are simulated with a prescribed probability q_{i1} of a species being recorded in each secondary unit of primary unit i at the start of the survey, q thus being constant within, but varying between, primary units, since there is potentially great variation in the percentage of ground covered by a species in the different primary units in which it is found. Specifically, logit(q_{i1}) is assumed to be normally distributed with mean μ, and we shall consider separately analyses based upon three values for μ. For each μ the analyses are performed with two different values for the variance of this underlying normal distribution, such that the coefficient of variation is (i) 0.5 and (ii) 0.8. Generally, although the species for which the survey is proposed are widespread the probability of occupancy in a secondary unit will still be low; we thus adopt three values for the mean (μ= -1, -2 or -3) which, with the two values of the variance associated with each, gives a total of six initial distributions of the occupancy probabilities for a species. These are considered to constitute a realistic range of real percentage occupancy measures for reasonably widespread species.

These values have been chosen to cover a reasonable range of likely values, comparable with a range of species' data in the Common Plants Survey (CPS), for example. Few species will be more common than is implied by μ=-1, for example (which corresponds to a primary unit in which just over 25% of potentially-selected secondary units are occupied), and species rarer than implied by μ= -3 (just below 5% of secondary units occupied) are unlikely to produce much power to identify all but the most catastrophic declines. Further illustration of the ranges of these starting populations are provided in Figures 8.2 and 8.3; with μ=-3 and a standard deviation of 1.5, the initial probability of a secondary unit being occupied is below 10% in around 70% of the primary squares, though with μ= -1 and standard deviation of 0.5, a more widespread species, 40% of squares have an initial occupancy probability of 20-30%. Figure 8.3 then shows how, when the variance is increased, the probabilities of secondary units being occupied show a greater 'spread', with most noticeably more of the high probabilities at the right-hand end of the figure.

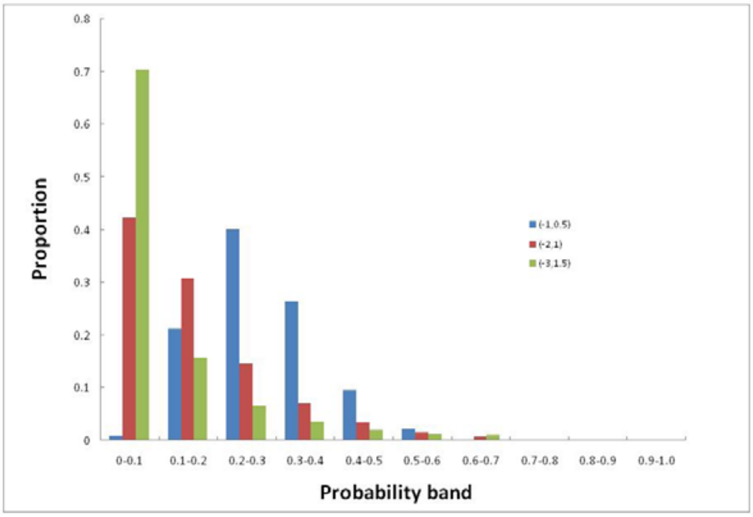


Figure 8.2. Distribution of site-specific probabilities of a plant species' presence in secondary squares at the start of the simulated data series. Three distributions are presented differing in the mean logit abundance (μ= -1, -2 or -3) and the standard deviation (σ= 0.5, 1 and 1.5 respectively). Diagram derived from frequency distributions of 1,000,000 randomly generated normal variates, logit transformed, in each case, the proportion (*Y*) of these transformed variates falling within the designated probability bands (*X*).

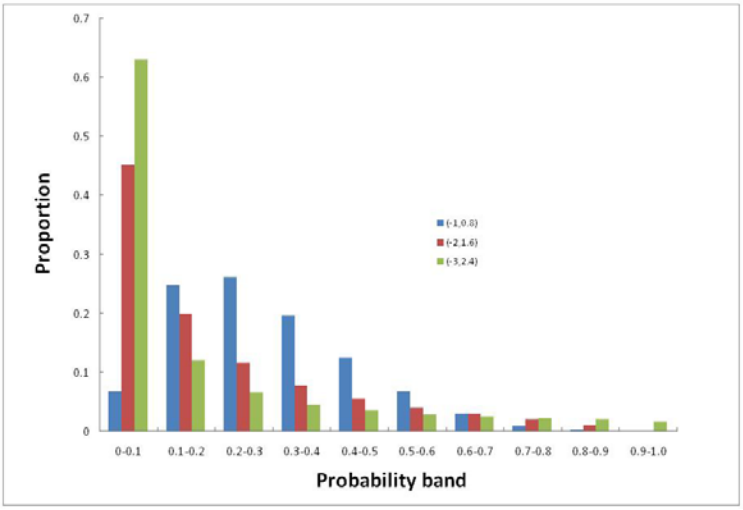


Figure 8.3. As in Figure 8.2 but with greater variance in the initial probabilities of a species being present; means and standard deviations are (-1,0.8), (-2,1.6) and (-3,2.4).

The second step is to define the relationship between occupancy at the secondary site level and occupancy at the primary site level using the parameters described above. We thus define the frequency, f_{it}, at a primary site *i* in year *t* as the number of secondary units recording the species, where in this case f_{it} ~ Bin(5,q_{it}) i.e. f_{it} is a binomial variable with 5 'trials' (secondary samples) and a probability q_{it} of being present in each sample. Further, if secondary sites are independent the probability that a species is entirely unrecorded at a primary site is given by the probability that it is not recorded at any of the secondary units:

$(1-q_{it})^5$

and thus we define the presence P_{it} at a primary unit (i.e. the species is recorded in at least one of the sampled secondary units) as a Bernoulli variable with probability of one minus the probability that it is not recorded in any of the secondary units:

$P_{it} \sim$ Bernoulli $(1-(1-q_{it})^5)$

Such initial occupancy probabilities are generated for each of the 100 primary units, and we then assume that the probability of occupancy decreases at all sites over a period of ten years. If q_{it} is the probability of a species' presence in a secondary unit within (primary) site *i* in year *t*, then we assume

$$\log it(q_{it}) = q_{1t} + a \times (t-1)$$

That is, the slope parameter *a* represents the annual change in the log odds ratio, consistent with the conventional approach to modelling binomially distributed data. Two values of *a* are used in simulating the data for power analyses, *a*=-0.011638 (corresponding to a 10% decrease in the odds ratio, the probability of occurrence divided by that of non-occurrence, over a ten-year period) and *a*=-0.02449 (a 20% decrease). On the real scale, these equate to a species which had a 50% probability of occupancy in the first year having a 47.5% probability of occupancy after 10 years of a shallow decline, and a 45% probability of occupancy after 10 years of a steep decline. For a species with 50% occupancy, they are therefore roughly equivalent to 5% and 10% population declines over the period. For a species with an initial occupancy probability of 10% these are equivalent to population declines of 9% and 18%, respectively. These two levels of decline are referred to henceforth as 'shallow' and 'steep' declines respectively for convenience.

Data were then simulated for the various combinations of the six specified values of initial abundance (occupancy probability) and two scenarios of subsequent decline. For each combination, data for 100 sites were simulated and independently replicated 500 times. In analysis, binomial models were fitted to each set of simulated data and the rate of decrease estimated from models fitted to the artificially generated observations of *f* and *P* in turn, with 'site' as a factor and 'year' as a continuous variable. The comparison in power between results based on modelling *f* and *P* provide an assessment of the advantages of using measures of occupancy at secondary sites within the primary units (power (frequency)), to those based solely on occupancy in the primary unit (power (presence)). The resulting estimates of the year coefficient $\hat{a}$ and their associated standard errors were stored; and the proportion of these estimates with 95% confidence limits not encompassing zero was then used as an estimate of the power of the model to identify the decline. These proportions are provided in Tables 8.1, and in Table 8.2 for the case of 10 secondary units per primary square.

The extent to which power is increased by accounting for the proportion of secondary units occupied, rather than merely using presence or absence at the primary level, is apparent. Power never approaches 50% in any of the analyses based upon modelling *P.* Further, power predictably increases with the number of secondary units within each primary unit.

Table 8.1. Estimates of power (% significant results from 500 replicates) in 12 different scenarios, defined by the initial abundance (μ) and the slope of the decline. Power is estimated from data in frequency form, and from simple presence in the record at the primary scale. Initial and final occupancies are derived from $\mu \pm 1.96(\mu \times CV)$, logit transformed – the range within which 95% of the probabilities of site-occupancy occur at the start and end of a ten year series. For example, for μ= -1 and a CV of 0.5, secondary squares within 95% of primary squares have probabilities between 12.1 and 49.5% of being initially recorded in a random secondary unit; this range declines to between 11.1 and 46.9% by the end of the survey.

μ	Initial occupancy (%)	Decline (of log odds)	Final occupancy (%)	Power % (frequency)	Power % (presence)
(a)	Coefficient of Variation = 0.5; 5 secondary sampling units				
-1	12.1, 49.5	10%	11.1, 46.9	17.6	12.2
-2	1.9, 49.0	10%	1.6, 46.4	13.0	12.4
-3	0.3, 48.5	10%	0.2, 45.9	9.0	8.8
-1	12.1, 49.5	20%	10.0, 44.0	57.8	34.6
-2	1.9, 49.0	20%	1.5, 43.5	37.0	24.6
-3	0.3, 48.5	20%	0.2, 43.0	24.0	17.8
(b)	Coefficient of Variation = 0.8; 5 secondary sampling units				
-1	7.0, 63.8	10%	6.5, 61.4	19.6	12.4
-2	0.5, 75.7	10%	0.5, 73.7	12.8	8.2
-3	0.0, 84.6	10%	0.0, 83.2	9.6	6.4
-1	7.0, 63.8	20%	5.8, 58.6	53.4	30.2
-2	0.5, 75.7	20%	0.5, 71.4	38.6	24.4
-3	0.0, 84.6	20%	0.0, 81.5	27.4	17.6

Table 8.2. Estimates of power (% significant results from 500 replicates) in 12 different scenarios, defined by the initial abundance (μ) and the slope of the decline. As Table 1, but based upon sampling 10 secondary units at each primary unit.

μ	**Initial occupancy (%)**	**Decline (of log odds)**	**Final occupancy (%)**	**Power % (frequency)**	**Power % (presence)**
(a)	Coefficient of Variation = 0.5; 10 secondary sampling units				
-1	12.1, 49.5	10%	11.1, 46.9	32.0	10.8
-2	1.9, 49.0	10%	1.6, 46.4	23.4	12.8
-3	0.3, 48.5	10%	0.2, 45.9	15.0	9.8
-1	12.1, 49.5	20%	10.0, 44.0	87.4	26.6
-2	1.9, 49.0	20%	1.5, 43.5	63.4	30.0
-3	0.3, 48.5	20%	0.2, 43.0	46.0	23.2
(b)	Coefficient of Variation = 0.8; 10 secondary sampling units				
-1	7.0, 63.8	10%	6.5, 61.4	30.0	13.6
-2	0.5, 75.7	10%	0.5, 73.7	23.8	10.4
-3	0.0, 84.6	10%	0.0, 83.2	13.6	9.4
-1	7.0, 63.8	20%	5.8, 58.6	85.0	24.4
-2	(0.5, 75.7)	20%	(0.5, 71.4)	71.0	26.8
-3	(0.0, 84.6)	20%	(0.0, 81.5)	47.2	18.2

With 10 secondary units power exceeds 80% at the top end of the abundance range considered here, and is nearly 50% for the rarest species in these scenarios. For a given μ (i.e. abundance), results are not greatly sensitive to the spread (determined by the standard deviation) of the initial distribution, at least not over the range considered here.

Power is, of course, dependent upon other features of the data including the duration and the number of primary sites sampled. Consider for instance the results of Table 8.2b, for the shallower decline (a=-0.011638). These are compared in Figure 8.4 with their counterparts from a study extended from 10 to 20 years, assuming that the same rate of decline continues throughout. Even with this shallow decline, power now exceeds 80% for any starting level of abundance and a frequency based model. The corresponding presence models likewise perform better but still fall short of 50% power. Of course, for conservation purposes, delayed identification of a decline severely limits attempts at remedial action.

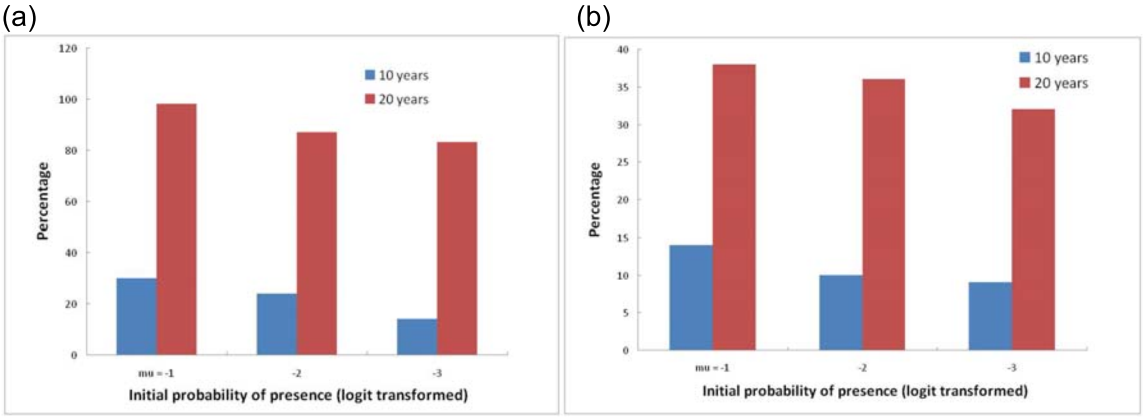


Figure 8.4. Improved power (*Y*, expressed as a percentage) over an extended survey: (a) frequency and (b) presence data for three values of μ (*X*).

The extent to which power is increased by raising the number of primary sites visited is demonstrated in Figures 8.5 and 8.6. Note that as all models contain an additional estimable parameter for each unique site, computation is very slow once samples reach a size of several hundred. The plots of Figures 8.5 and 8.6 are each therefore based on 500 repeated replicates. While this is comfortably sufficient to approximate the power to detect the declines indicated, the small sampling error remaining accounts for the slight deviations in the shape of these figures from those of smooth theoretical curves.

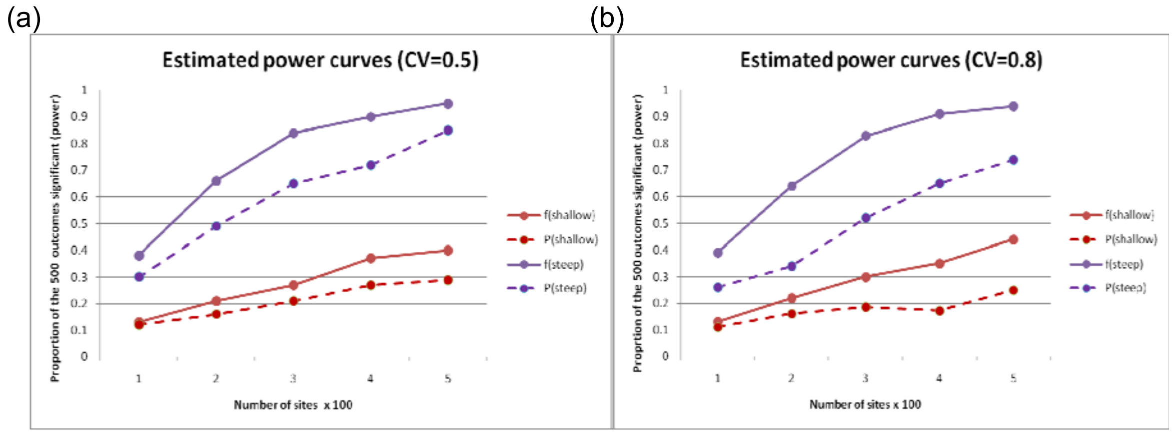
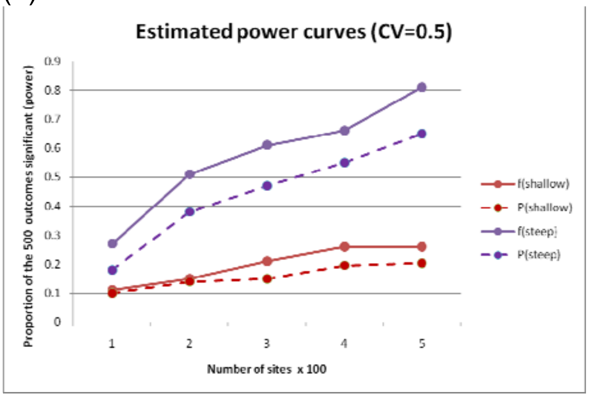


Figure 8.5. Power with sites increasing from 100 to 500: μ=-2, all other factors as Table 8.1.

It is seen that for μ= -2 the frequency-based model achieves an impressive 90% power to detect the steeper decline at between 300 and 400 sites (Figure 8.5), with even the analysis of data at the level of primary square presence reaching about 70% for 400 sites. The shallower decline remains poorly detected even at 500 sites. Even for the rarest species among these scenarios (μ= -3, Figure 8.6), the steeper decline is well identified by frequency data for a sample size of >300 sites.

(a)



(b)

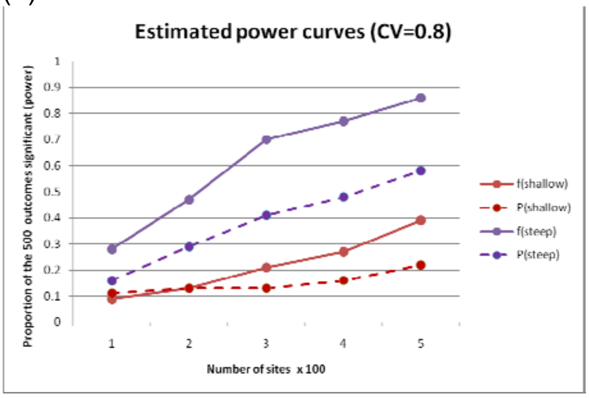


Figure 8.6. Power with sites increasing from 100 to 500: μ-3, all other factors as Table 8.1.

8.3 Alternative approaches to modelling range-change in widespread plants

We now consider the possibility of obtaining more robust inference on temporal trends in plant abundance by adopting alternative approaches to gathering, and modelling, data on a large scale.

8.3.1 A doubly-censored logit normal model for common plant abundance

Plant surveys often record percentage ground cover of species within one of a small number of quantitative, but broad categories. We define the recorded proportional cover of a species as Q, $0<Q<1$, and assign data collected under the Common Plant Survey (CPS) to the following categories:

1:	0	<Q<	0.01
2:	0.01	<Q<	0.05
3:	0.05	<Q<	0.25
4:	0.25	<Q<	0.5
5:	0.5	<Q<	0.75
6:	0.75	<Q<	1

We discuss here a new method of estimating percentage cover of a species from such data, with appropriate measures of statistical precision, and the use of such models to track annual variation in abundance. Two further categories of data are available; records of simple presence (p) or absence (a) at sites. For computational reasons we categorise these records as follows:

p: $\varepsilon < Q < 1$
a: $0 < Q < \varepsilon$

That is, we define a value, ε, close to zero, such that for values of Q above ε we assume presence, and below which we assume absence. We have not found results to be sensitive to the exact value of the arbitrarily small threshold ε – here we take ε= 0.000001.

The true value of Q at a visit is unknown, and assumed to be normally distributed on a logit scale. We propose here a model in which X = logit(Q) = log(Q/1-Q) ~ N(μ,σ^2).

Thus, we have an inequality $u<Q<v \Leftrightarrow$ logit(u)<logit(Q)<logit(v). If we denote by L_i and U_i the logit-transformed lower and upper bounds of coverage category *i*, then for each coverage category (*p*,*a*,*1*,*2*,*3*,*4*,*5*,*6*), we have the probability P(*i*) that an observation lies in the category *i*:

$$P(i) = \Phi(U_i) - \Phi(L_i)$$

where Φ(.) is the cumulative distribution function of a normal distribution with mean μ and variance σ^2.

It is therefore possible to compute a likelihood function for a set of CPS data under this model, which can be maximised via numerical methods to estimate the parameters μ and σ. In this report we provide illustrations of the fitting of such a model to both simulated data and a set of observations from the CPS database. Bespoke code written in R was adopted for the purposes of model fitting on the logit scale. The parameter estimates $\hat{\mu}, \hat{\sigma}^2$ were then converted into expected values for percentage coverage, a more intuitively interpretable quantity, by then generating 100,000 random numbers X' $\sim N(\hat{\mu}, \hat{\sigma}^2)$ and taking the sample mean of the corresponding back-transformed Q'.

It is then possible to develop the model such that the mean percentage cover on the logit scale μ is a function of time (producing an estimate of average annual change) or indeed any other environmental covariate(s) of interest. Here we shall consider simple logit-linear time trends only: μ = α + β×year, but note that the method is potentially also useful in identifying, for example, climatic factors defining the range of a species. Standard likelihood-ratio tests can then be applied to test the hypothesis β=0, i.e. that of no long-term change in abundance.

The logit-linear time model is illustrated using CPS data on four species in Figures 8.7 (parameters presented on the logit scale, on which the model fitting is carried out) and 8.8 (where the same results are converted to percentage cover). Note that three categories of survey 'unit' are represented in the CPS; the analyses in these figures are based upon those from the 'linear' plots only. Figure 8.7, and the symbols in Figure 8.8, represent simple estimates independently obtained for each year 2000-2007, excluding 2001 (the year of foot-and-mouth disease). The smooth lines represent the fit of the logit-linear trend to all years' data simultaneously. Note that confidence limits are given for the estimates on the logit scale; similar confidence limits for the estimated percentages demands considerably greater computation time and are not yet computed.

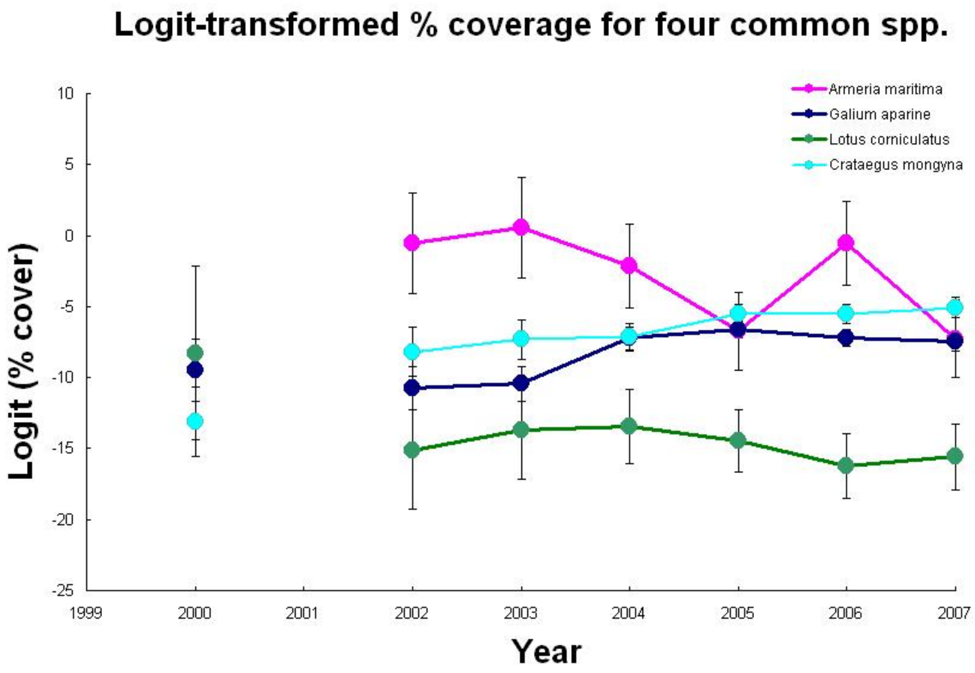


Figure 8.7. Estimated logit-transformed percentage cover for four common species, 2000-2007. Bars indicate estimated asymptotic 95% confidence limits.

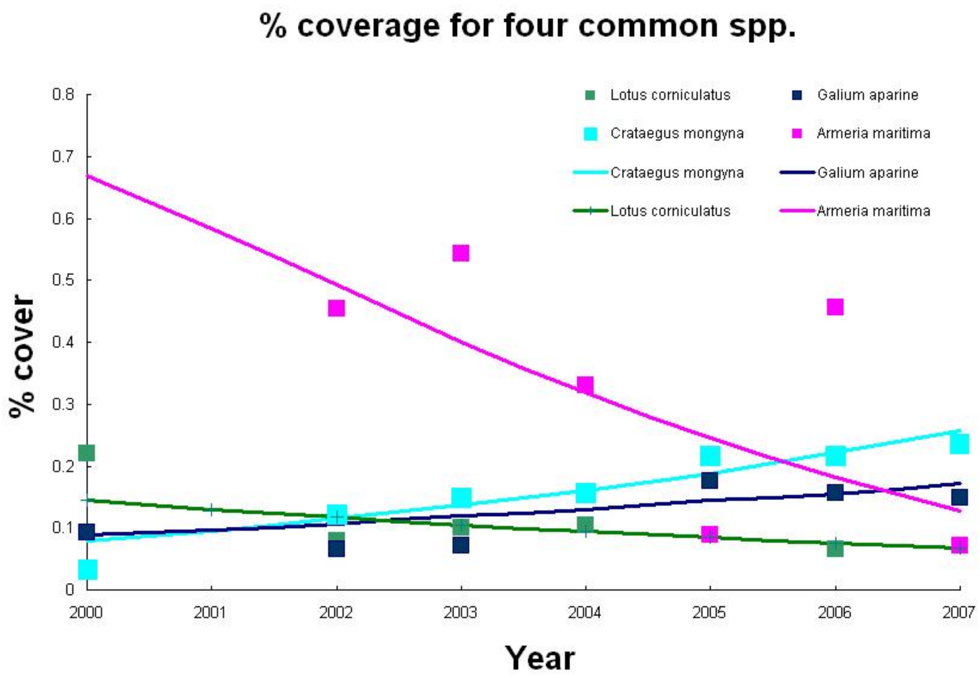


Figure 8.8. Estimated percentage cover for four common species, 2000-2007. Lines show constant annual rate of change on logit scale.

Tests of significance are given in Table 8.3, and these indicate significant change through time in all species except *Armeria maritima*, with *Lotus corniculatus* being the only species to show a significant decline.

Table 8.3. Assessing the significance of temporal trends in four species via doubly-censored logit models: log-likelihood ratio statistics.

Species	Log-likelihoods under model		LRT statistic (* = significant, p<0.05)
	$\mu=\alpha+\beta \times$ time	$\mu=\alpha$	
Armeria maritima	61.59	63.09	3.00
Galium aparine	4753.40	4762.33	17.86*
Lotus corniculatus	912.15	914.73	5.16*
Crataegus monogyna	3558.94	3581.03	44.18*

A further investigation of model performance using simulated data is given in Table 8.4.

Table 8.4. Average standard errors for $\hat{\mu}$ across 1000 sets of data simulated with parameter values and sample sizes indicated, and fitted with the doubly-censored logit normal model.

μ	σ	100 sites	200 sites	300 sites
-10	10	1.61	1.08	0.87
-8		1.29	0.86	0.70
-6		1.04	0.71	0.57
-10	6	1.29	0.88	0.70
-8		0.89	0.60	0.48
-6		0.64	0.44	0.36

Estimates of mean logit-transformed abundance μ were obtained by generating 1000 replicates of a single-year study in which 100, 200 or 300 sites were visited, assuming that the percentage cover attributable to a hypothetical species was distributed according to one of the values chosen for μ and σ. For simplicity we assume all observations are made in categories 1-6 only, and none are represented as simple absence or presence. Average standard errors estimated from the 1000 replicates are presented, and illustrate the level of precision that can be anticipated for such data, whose distributions cover a range realistic for many species in the CPS data. As would be expected, precision is increased with greater sample size and with smaller variability in sites. An extended simulation study along these lines would be timely, to ascertain power to detect declines of specified magnitude (along the lines of Figure 8.8) within a certain period of time, and to investigate the extent to which the proposed method might improve on more conventional analyses of data which is 'presence/absence' or proportion of secondary samples with presence.

8.3.2 Site-occupancy models and ‘nested’ designs

We conclude by noting two other approaches to modelling the detection of species not considered in detail here, but with apparent potential for the surveillance of plants.

In considering ’presence/absence’ style data as conventional binomial variables, as in section 8.1, it is implicitly assumed that no species are missed during the survey. The validity of this assumption is likely to vary between species and to diminish with a larger sampling area. It is impossible at a single visit to distinguish between species genuinely absent from a site, and those present (possibly in small number) but merely overlooked. Further, the variation in detectability caused by flowering makes certainty slightly more difficult for particular species. The availability of repeated visits to a site, however, allows the fitting of a more sophisticated model in which the probabilities of a species being absent, or merely overlooked given that it is present, are separately estimable. The estimates of occupancy probabilities obtained are therefore unbiased by variation in the probability of detecting, or otherwise, a species that is present. The secondary sampling protocol described in this report makes this a viable option by, for example, using the repeated records in the sub-units of the primary sites to estimate without bias the probability of presence or absence at this primary level, even in the absence of additional temporal replication of visits within each season. Such methods modelling occupancy are now widely-used and theoretical development is both well-documented and continuing to advance (e.g. Mackenzie *et al* 2006).

A further method, appealing in its simplicity, is to adopt the ‘nested’ designs of Whitaker *et al* (Stohlgren *et al* 1995). The concept is conceptually easy to adopt in the field – a thorough record is made of species encountered in an initial location, from which the observer then progresses, recording only additional species encountered within a specified distance. However, the small scale, and complex set-up in Stohlgren *et al* (1995) suggest the practicalities may be more difficult. We note here the similarity with ‘Timed Species Counts’ previously adopted for censusing birds in tropical regions, in which encounters with birds are disregarded once the time of the first encounter with the species has been recorded. See Freeman, Pomeroy and Tushabe (2003) for an account of the increase in precision attainable in such a study compared to a simple ‘frequency’ model of the kind in Section 8 of this report, over an identical time period.

8.4 Conclusions on analytical power and sample size

Possibly the simplest design for a wide-scale plant monitoring scheme is the simple listing of species recorded at a large sample of sites. We have seen that power to detect change over a period of time is substantially increased by modifying the design to include sampling at two different levels and gathering species lists on the smaller, more practically realistic units. The natural statistical model is binomial in form, and fits simply into a binomial GLM or GLMM framework in any of a range of widely-used statistical software, as often used in ecology with presence/absence data.

Even for a widespread species, the probability of its being both present and recorded in a random square of manageable size is low; by means of Monte Carlo simulation we have shown however that over the range of abundances considered in this report declines over time in the odds of a species being located can be found with an acceptable level of >80% power for <500 primary sites containing five smaller secondary units. Studies of this size tend to give low power (<50%) if the data are analysed as presence/absence at the level of the primary unit. Power does of course depend upon a large number of factors; it may be increased from the levels given here by increasing the size of the sample, or the duration of the time series, and it will be higher for more abundant or more rapidly changing species, although it may also decrease if a substantial proportion of visits are not carried out, or if the fit of the underlying binomial model is poor. It is assumed here that species lists at the level of the secondary units are comprehensive, and species present are not missed. Separately estimating probabilities of presence and subsequent conditional probabilities of detection is the domain of site-occupancy modelling, and is discussed below.

Note that the estimates of percentage cover presented for example species in Figure 8.8 lie largely in the range 0-20%; we have found this to be typical of the full set of species for which the CPS contains appreciable data. Although this of course says nothing of the extent to which a species is clumped or dispersed within a square, these figures are similar in range to the recording probabilities q_{it} we adopted in the simulation study.

Though considered in less detail, we have also discussed a number of alternative, non-standard models. The field protocol is more complex, as is the method of model-fitting, but each of them has potential to increase the efficiency of a monitoring programme and is, we believe, worthy of further investigation. The adoption of nested or site-occupancy models is now well-established in other contexts. We have also provided a preliminary investigation of an original approach, based on the collection of data in 'bands' of percentage cover and the adoption of a doubly-censored logit-normal model. Initial results here are promising but should be viewed as exploratory; the implicit assumption of independence between identically-distributed observations that we have made here to facilitate fitting is certainly violated, by, for example, a degree of serial correlation between series of data from any individual site. Accounting for such refinements complicates the fitting considerably more, but this too we believe is an approach meriting further investigation.

Two other statistical issues not dealt with above require further consideration during the design of any new scheme:

Geographic stratification: Given the proposed stratification of primary units within vice-counties we will need to determine the minimum number of 1km squares required per Vice-county based on total population density. Similarly, if the reporting of environmental changes is required at the regional level we will need to understand how aggregation of vice-counties and their associated sample sizes would impact on this. Consideration will also be needed to determine whether more samples will be needed to make country (Scotland and Wales) assessments statistically robust and capable of illustrating national trends (e.g. more sample points for Countryside Survey in Wales).

Bias in sampling: Arranging access to private land is known to be an obstacle to participation. One solution is to survey from public rights of way alone. Although this can introduce bias from a statistical perspective it is less of a concern than reducing sample size through reduced participation. It should also be noted that any bias might actually be quite small. For example analyses of broad habitat classes for the BTO's Breeding Birds Survey, (which is mostly done from rights of way) did not detect any significant effect or bias. Of the Methods presented in this report only Method 2 and possibly Method 7 are likely to be immune from this bias as the positioning of plots or transects under Methods 3-5 are likely to be influenced by access routes.

8.5 Recommendations

i Comprehensive species lists at a large scale (say 1 x1 km) cannot be reliably ascertained. Sub-sampling a number of smaller secondary units within can be achieved in greater detail. The adoption of binomial models for the ‘frequency’ of detections of a species in secondary units within each primary unit increases power, with respect to a simple modelling of whether a species was recorded in primary units.

ii At least for reasonably widespread species, reasonable power can be achieved to detect relatively modest declines with a practical number of primary units (i.e. volunteers). Though power can be increased by increasing the number of secondary units, there is of course a trade-off between this and the number that a volunteer can reasonably accomplish within a single day.

iii A number of alternative approaches are discussed. All retain a relatively straightforward field protocol, yet have various advantages. Some model percentage cover rather than mere presence, hence have greater sensitivity to change. Others open up the possibility of relaxing the assumption that species are not missed where present, and hence correct for any bias in violating this assumption. An invaluable resource is available in the form of large-scale, long-term plant survey data and this has perhaps yet to be utilised to its full potential. Statistical research into the development of models such as these is continuing apace and we believe more research into the development of bespoke models for UK plant trends could enormously increase our ability to monitor these important components of the ecosystem.

9 Implementation of the scheme

Whilst the parameters of this project sought the design of an 'ideal' plant surveillance scheme, the theoretical ideal scheme may not prove practical to deliver, especially given the constraints of the scheme needing to be low cost i.e. implemented largely through volunteers and limited use of professionals in the field. The project team have therefore sought to take factors effecting implementation, and the use of a volunteer network, into account in the proposed methods outlined.

The ultimate success of a surveillance scheme aiming to detect changes in common species throughout a wide range of environmental drivers relies on maximising data collation to provide robust analysis for the determination of trends. Subsequently, a surveillance scheme must be sufficiently engaging to bring in participants and support long-term commitment. The project team has been able to draw on its extensive experience in the management and implementation of surveillance schemes carried out by volunteer recorders.

It must be stressed that facilitation of a cost-effective and sustainable surveillance scheme, which makes the best use of volunteers, requires adequate project management, promotional activity and volunteer support mechanisms. Additionally there is the need for data capture, statistical analysis and interpretation. Appropriate resourcing is essential for delivery.

9.1 Volunteer engagement

It is necessary to understand the reasons why people may choose to volunteer on a plant surveillance scheme. These can include a passion for plants, enjoyment of walking, wanting to see how the countryside is changing, learning more about the environment, contributing to an organised survey/conservation programme, learning more about plants, a reason to go out, a new hobby, as a contribution to mental and physical health, and for professional development.

There are also barriers and constraints to volunteer participation - the following issues have arisen from volunteers engaged in Plantlife's Common Plants Survey:

- Variation in skill level: from enthusiastic novices to experienced botanists. All levels need to be catered for (such as through the use of manageable target lists of species to record) and encouraged in order to maximise participation.
- Isolation: whilst some volunteers are happy to survey alone, Health and Safety guidance suggests surveying in pairs. Surveying as part of a group is a popular option.
- Survey method: must not be too complicated, time-consuming or boring. For example, within the Common Plants Survey the assessment of abundance according to a 6 category scale was considered too difficult by many.
- Location of survey: most volunteers prefer to survey close to their home but some are prepared to travel distances. The cost of transport may be an issue for some volunteers.
- Land ownership and access permission: finding ownership and gaining permission to survey is a major barrier for some volunteers
- Plots: can be difficult for surveyors to locate in the absence of relevant ground features or GPS and therefore is off-putting
- Species presence: the absence of target species within recording plots can be a barrier to participant retention. Varied activities can help to overcome this.

- Training: training events facilitate participation and volunteer retention. Workshops structured to the needs of the surveyors should be accessible by public transport and locally based.
- Feedback: Surveyors appreciate feedback including prompt replies to email enquiries, regular newsletters or yearly reports.

In relation to the suggested use of 500 primary samples (monads) this can be compared to 450 CPS squares (monads) surveyed in the 2009 field season and 700+ tetrads covered by BSBI members in Local Change (2003/04 latest round).

9.2 Project delivery

The support and delivery frameworks for a plant surveillance scheme will need careful consideration and comparisons with existing surveillance schemes (such as Breeding Bird Survey and Butterfly Monitoring Scheme) should be made so that best practise is adopted.

Aspects which require consideration include: project management and the benefits of partnership working; central and local support structures for volunteers; technology in the field; training opportunities; easing access issues; data capture; and statistical analysis requirements.

The costs associated with the plant surveillance scheme should compare the use of volunteers only in the field as well as a combination of volunteers and professionals. The extent to which payment will be needed for field work should depend upon issues such as ease of access to sites and species/habitat complexity.

10 Recommendations and field testing

10.1 Target species

- Participants should have the option to record either all species they encounter (Level 3), if they are sufficiently experienced to do so, or from one of the two subsets of target species: 30-40 easy to moderately easy target species from each habitat type (Level 2) or 20 easy to identify species from each habitat type (Level 1). All Level 1 species to be included in the list for Level 2;
- Target species should not be selected on the basis of traits although the traits profile of those selected needs to be checked to ensure they are broadly representative of the flora as a whole;
- With the exception of a few restricted habitats (e.g. montane) target species should occur in greater than 750 hectads in Britain (based on data from Preston *et al* 2002) and be easy or moderately-easy to identify;
- 20-30 target species should be selected for each broad habitat (if EUNIS Level II are used then the numbers may need to be refined due to the greater number of habitats covered);
- Targets should include a subset of 'specialists' e.g. restricted to a single Broad Habitat or EUNIS Level II habitat, widespread axiophyte, etc;
- Targets should include a subset of taxa which are important food resources for other taxa (e.g. birds, pollinators);
- If further target species are needed then these should be selected randomly from those remaining;
- Field testing will be required following the selection of target species to ensure that there is an adequate encounter rate.

10.2 Sample stratification

- Primary units (sample) to be 1km squares;
- Stratification of primary units to be based on Watsonian vice-counties;
- The number of primary units per strata should be weighted to the number of recorders present in each county, using human population size as a proxy measure;
- Define a minimum and maximum number of samples per strata to reduce geographic bias in sample coverage;
- Primary units should be selected at random from the 'pool' of random 1km squares currently being monitored as part of existing surveillance schemes, including other taxonomic groups, plus non-random squares included in the BSBI's Local Change network (SW corner 1km square of each tetrad);
- No field testing required but simulations would be useful prior to survey to estimate the numbers of squares that are likely to be allocated (i.e. recorder effort needed).

10.3 Survey method

- We recommend that one of the following methods is adopted following field testing, further statistical analyses and simulations:

    - 100m grid cells stratified by habitat (Method 3);
    - Walk/transect plus habitat plots (Method 5b);

- Nested plots (Method 7).

Field testing of these methods should focus on:

- Time taken to complete full sampling including locating 1km squares/plots in a range of landscape types and with 5 or 10 secondary sample units;
- Access issues, methods to overcome these and differences in results using public access only versus full access;
- Whether the frequency of recorded species within each plot/nest could be estimated in the time available as well as presence;
- Practicalities of surveyors identifying main habitat types from a predefined list;
- Practicalities of surveyors selecting ‘representative’ habitat patches (Methods 3 & 5b);
- Design where patches are smaller than the plot size (e.g. streams, hedgerows, flushes, etc.) (Method 3 & 5b);
- Size and position of nests (Method 7).

10.4 Optimal time to survey

- Assuming a single visit, late June/July would be the optimal time to survey the majority of British habitats;
- Exceptions include Broad-leaved Woodland (June) and Saltmarsh (August);
- Three visits would be ideal with visits in the months either side of the optima;
- Field testing should assess the differences in recording efficiency based on 1, 2 or 3 field visits across a range of habitat types.

10.5 Parameters to record

- Precisely which additional attributes (beyond species presence) can be collected will depend on the method of survey chosen, in particular the size of plots/nests as many parameters are difficult to record above 2m;
- It is recommended that the collection of attributes in Table 7.1 are tested in the field with each survey method;
- This should assess:
    - how easy the data are to record?
    - the amount of time they add to the survey time?
    - value of the data collected?
    - variability in data collected between recorders?

10.6 Analytical power and sample size

- The power to detect change over time is substantially increased by recording more secondary plots within each of the primary sampling units, and gathering species lists on these smaller, more practically viable units;
- Even for a widespread species, the probability of its being recorded in a random square of manageable size is low;

- Over the range of abundances considered substantial declines over time in the odds of a species being located can be found with an acceptable level of >80% power for <500 primary sites containing five smaller secondary units;
- Studies of this size tend to give low power (<50%) if the data analysed are instead degraded into mere presence/absence at the level of the primary unit. It may be increased by increasing the size or duration of the sample, and it will be higher for more abundant or more rapidly declining species (although it may also decrease if a substantial proportion of visits are not carried out or if the fit of the underlying binomial model is poor);
- A number of alternative, non-standard models are available (such as nested or site-occupancy models) and each has the potential to increase the efficiency of a monitoring programme and is, we believe, worthy of further investigation.

10.7 Implementation

- The support structures required for the plant surveillance scheme should be considered;
- Comparisons with existing surveillance schemes should be made so that best practice is adopted;
- Aspects which require consideration include: project management and the benefits of partnership working; central and local support structures for volunteers; technology in the field; training opportunities; easing access issues; data capture; and statistical analysis requirements.

11 References

BRAITHWAITE, M.E., ELLIS, R.W. & PRESTON, C.D. 2006. *Change in the British Flora 1987-2004*. Botanical Society of the British Isles, London.

COPE, T. & GRAY, A. 2009. *Grasses of the British Isles*. Botanical Society of the British Isles Handbook No. 13. BSBI, London.

CURRALL, J.E.P. 1987. A transformation of the Domin scale. *Vegetatio* **72**: 81-87.

FITTER, A.H. & PEAT , H.J. 1994. The Ecological Flora Database. *Journal of Ecology* **82**: 415-425.

FREEMAN, S.N., POMEROY, D.E. & TUSHABE, H. 2003. On the use of Timed Species Counts to estimate avian abundance indices in species-rich communities. *African Journal of Ecology* **41**: 337-348.

GRIME, J.P., HODGSON, J.G. & HUNT, R. 2007. *Comparative Plant Ecology: a Functional Approach to Common British Species*. Second edition. Castlepoint Press, Dalbeatie.

HILL, M.O., PRESTON, C.D. & ROY, D.B. 2004. *PLANTATT. Attributes of British and Irish Plants: Status, Size, Life history, Geography and Habitats*. NERC Centre for Ecology and Hydrology, Monks Wood.

HILL, M.O., PRESTON, C.D., BOSANQUET, S.D.S. & ROY, D.B. 2007. *BRYOATT. Attributes of British and Irish Mosses, Liverworts & Hornworts*. NERC Centre for Ecology and Hydrology & Countryside Council for Wales.

JACKSON, D.L. 2000. *Guidance on the interpretation of Biodiversity Broad Habitat Classification (terrestrial and freshwater types): Definitions and relationship with other habitat classifications*. *JNCC Report* No. 307. JNCC, Peterborough.

JERMY, A.C., SIMPSON, D.A., FOLEY, M.J.Y. & PORTER, M.S. 2007. *Sedges of the British Isles.* Botanical Society of the British Isles Handbook No 1. Third edition. BSBI, London.

JNCC. 2009. *The UK Terrestrial Biodiversity Surveillance Strategy*. Version February 2009. JNCC, Peterborough.

KEITH, S.A., NEWTON, A.C., MORECROFT, M.D., BEALEY, C.E. & BULLOCK, J.M. 2009. Taxonomic homogenization of woodland plant communities over 70 years. *Proceedings of the Royal Society B* **276**: 3539-3544.

LOCKTON, A. 2006. Coordinator's corner. *BSBI News* **103**: 57-58.

MACKENZIE, D.I., NICHOLS, J.D., ROYLE, J.A., POLLOCK, K.H., BAILEY, L.L. & HINES, J.E. 2006. *Occupancy estimation and modelling*. Elsevier.

MASKELL, L.C., NORTON, L.R., SMART, S.M., SCOTT, R., CAREY, P.D., MURPHY, J., CHAMBERLAIN, P.M., WOOD, C.M., BUNCE, R.G.H. & BARR, C.J. 2007. *CS Technical*

Report No.2/07. Vegetation Plots Handbook. Centre for Ecology and Hydrology, Natural Environment Research Council.

MORECROFT, M.D., BEALEY, C.E., BEAUMONT, D.A., BENHAM, S., BROOKS, D.R., BURT, T.P., CRITCHLEY, C.R.N., DICK, J., LITTLEWOOD, N.A., MONTEITH, D.T., SCOTT, W.A., SMITH, R.I., WALMSLEY, C. & WATSON, H. 2009. Review: the UK Environmental Change Network: Emerging trends in the composition of plant and animal communities and the physical environment. *Biological Conservation* **142**: 2814–2832.

PRESTON, C.D., PEARMAN, D.A. & DINES, T.D. 2002. *New Atlas of the British and Irish flora*. Oxford University Press, Oxford.

PAGE, C.N. 1997. *The Ferns of Britain and Ireland*. Second edition. Cambridge University Press, Cambridge.

SMART, S.M., BUNCE, R.G.H., MARRS, R., LE DUC, M., FIRBANK, L.G., MASKELL, L.C., SCOTT, W.A., THOMPSON, K. & WALKER, K.J. 2005. Large-scale changes in the abundance of common higher plant species across Britain between 1978, 1990 and 1998 as a consequence of human activity: tests of hypothesised trait representation. *Biological Conservation* **124**: 355-371.

STACE, C. 2010. *New flora of the British Isles*. Third edition. Cambridge University Press, Cambridge.

STEWART, K.E.J., BOURNE, N. & THOMAS, J. 2001. An evaluation of three quick methods commonly used to assess sward height in ecology. *Journal of Applied Ecology* **38**: 1148-1154.

STREETER, D., HART-DAVIES, C., HARDCASTLE, A., COLE, F. & HARPER, L. 2009. *Collins Flower Guide*. Harper Collins, London.

STOHLGREN, T.J., FALKNER, M.B. & SCHELL, L.D. 1995. A Modified-Whittaker nested vegetation sampling method. *Vegetatio* **117**: 113-121.

WALKER, K.J., PRESTON, C.D. & BOON, C. 2009. Fifty years of change in an area of intensive agriculture: plant trait responses to habitat modification and conservation, Bedfordshire, England. *Biodiversity & Conservation* **18**: 3597-3613.

Annex 1

Notes from the plant surveillance discussion group, January 2010

Incorporating comments from Quentin Groom, Mick Crawley, Bob Ellis, Alex Lockton, Chris Preston and Simon Smart.

What are the objectives of 'ideal' surveillance?

- We need to be absolutely clear about the questions we are trying to answer and how we are going to analyze the data;
- We need to make a clear distinction as to whether we will measure changes in abundance and/or distribution as both require very different methods;
- The most realistic aim (given volunteer constraints) would be to record large-scale changes in the distribution of species in relation to major pressures (changes in soil nutrient status, land management intensity, habitat condition and the effects of policy mechanisms to enhance overall biodiversity) – we currently do this well and it would appeal to many recorders;
- More challenging would be to measure fine-scale changes in the abundance of plant species – this will only appeal to a minority of recorders (see below);
- Regardless of approach, geographic/habitat stratification is needed (see below) as it is important to know what is happening in the whole landscape, not just in small landscape relics.

Who will do the monitoring and what challenges does this pose?

- Volunteers, from beginners to experts – project therefore needs to be practical and appealing;
- Most volunteers will do distribution recording – only a small minority would be willing record permanent quadrats in remote and possibly rather dull places;
- Recorder expertise will be a major factor affecting the results of any scheme;
- Involving all levels of botanist will be more challenging - requires a hierarchy of things to record which would be difficult to present clearly and relies on people knowing how good they are.

Advantages of selecting target species?

- Allows participation of less-skilled botanists;
- Reduces recorder error.

Disadvantages of selecting target species?

- Assumes that there is an optimal set of target species and that we could predict (we would be bound to select the wrong species);
- Compromises the completeness of surveillance and its ability to detect unanticipated changes;

- Biases the data with our preconceptions of change;
- Inefficient sampling; better to record everything (this is built into the psyche of most recorders) – we can always focus on a subset of targets in the analyses.

Should recording focus on changes in abundance or distribution?

- Plants are changing in small and large scale spatial distribution within the UK; this recording would appeal to volunteers (and it does not require randomization, but it will benefit from stratification);
- Volunteers could potentially record the most dramatic changes in relative abundance that occur within geographic ranges but, since this involves the consistent and repeated estimation of plant abundance, it will be less appealing;
- Might require a much smaller, dedicated series of plots.

What are the limitations to recording changes in abundance?

- Plants are changing in abundance within their existing ranges: assessing these changes requires stratified random and replicated permanent quadrats;
- Abundance is difficult to record and changes through the year, between recorders, etc – ordinal categories therefore needed;
- The botanical recording community are not well-equipped to do this;
- This work would never appeal to more than a tiny minority of our members.

Important things to consider in recording ‘species’?

Only record species or species aggregates for the likes of *Taraxacum* and *Rubus.*

- Anything unidentified should still be noted down with its genus/family;
- If using targets phylogenetic autocorrelation is an issue as closely related indicator species will tend to behave more similarly in comparison to more distantly related species.

Are plants/plant traits useful indicators of environmental change?

- Plants are poor indicators; they respond slowly and in unpredictable ways (e.g. monitoring soil nitrogen content directly might be a better measure of eutrophication);
- Traits are of unproven usefulness. Assembly rules based on traits have failed to produce convincing insights (we do not know why coexisting species coexist). Loss of ecosystem function as a result of non-random or random extinction of species is a very hot topic, but not one to which volunteer recording is likely to contribute in a major way.

Which plant traits are the best indicators of environmental change?

- Ellenberg numbers provide the most useful means of ensuring good coverage of plant traits e.g. soil pH & nitrogen, atmospheric pollution input, Raunkaier’s life forms (trees, herbaceous perennials, annuals, geophytes, etc.);
- Annuals respond in different ways (i.e. change will be inversely proportional to plant longevity);

- Other summary variables include woody cover, grass:forb ratio, graminoid:forb ratio, vascular plant:Sphagnum ratio on bogs and heaths, Grime's C-S-R?

What are the advantages of using small permanent plots (non-random, random or stratified-random)?

- Small plots have a handful of species which are mainly habitat dominants;
- Change in habitat dominants would be most consequential; change in rare species could be argued to be functionally trivial;
- Allows you to select habitat dominants in an unbiased way;
- By repeating measurements on the same plot you can much more easily account for the variance of the data (real change versus sampling variation);
- The reason we now use non-linear, unitless expressions of change is because we don't have any permanent baseline survey we can rely on.

What are the disadvantages of using small plots?

- Potentially complex: different habitats need different shapes, sizes and sampling frequency;
- The large number of plots needed;
- Unappealing to volunteers (see above);
- Miss changes obvious changes in distribution easily detected via routine botanical recording (e.g. spread of *Cochlearia danica* on roadsides).

How would samples be stratified?

- Samples should be stratified by geographic location (e.g. tetrads in 100km squares) and by habitats within tetrads (habitat lists are much more useful than tetrad or 1km square lists);
- Habitat strata to include Broad Habitats or NVC (plus any missing habitats)?;
- NVC allows all plant communities to be properly represented although there are two main;
- problems: (a) subjectivity / unfamiliarity / lack of expertise in identifying types, (b) under-sampling of ordinary vegetation types / ecotones / gradients;
- Broad Habitats provide a much simpler system but we would need to devise robust subtypes within each category (woodland, grassland, mire, heath, aquatic, maritime, open, etc).

Should some Broad Habitat/NVC types be ignored (e.g. urban, standing water, bracken, etc.)?

- No. Urban, standing water, bracken are some of the habitats that exhibit the highest rates of change.

Do we need random samples to detect change?

- Recorders should not be choosing the plots;

- Truly random plots should be avoided – the these would be endless repetitions of the ordinary;
- Sampling should therefore be random within strata (see above);
- Four replicates per strata is a reasonable compromise between statistical power and time;
- Commitment.

What are the key environmental drivers?

- Land use change;
- Changes in the type of crops and trees grown;
- Changing in stocking densities and the type of animals reared;
- Climate change;
- Changes in salting of the roads;
- Invasive plants.

What is the most appropriate recording interval?

- The recording interval should be one year for permanent quadrats (so that year effects can be separated from trends) and 5 years for geographic surveys;
- Once a year maybe too much for very sensitive vegetation in wetlands.

What issues do we need to consider for analyses?

- Any analysis needs to avoid the problem of spatial autocorrelation;
- We can model the spatial covariance in the presence of fixed effects and other kinds of random effects using generalized least squares;
- Skill of the recorder.

Proposed method of occurrence in nested tetrad design

- Stratified random sample or a systematic grid of tetrads (former preferred);
- List all species (or indicator-dominants) in a random sample of small plots (within habitat strata?) then move out to census only additional species in a series of larger nests, ending with the whole tetrad;
- When doing this note the Broad Habitat within which first seen (plus additional?);
- Options could be tested for power and detectability of change among different species groups by simulating options based on data we already have.

Can we abandon an unbiased sample design?

- Given high cost of professional schemes careful design considerations could be reduced;
- This might give us an a rather biased but very well worked set of tetrads and nested plots that are recorded comprehensively and more frequently;
- Set-up a wide-ranging core set of ‘sentinel’ tetrads adopted by local volunteers where comprehensive census of all species is undertaken with maximum efficiency using a nested design from small to large plots;

- The data may end up having to be analysed and interpreted in ways that do not support unbiased inference to unsampled areas;
- The data could still be analysed along the lines of LC and their signals corroborated by comparison with other more unbiased schemes that nevertheless record far fewer species.

Annex 2

Correlations between Broad Habitats and EUNIS Level II habitats

BH code	Broad habitat name	EUNIS Level II	EUNIS habitat name
1	Broadleaved, mixed and yew woodland	E5	Woodland fringes & clearings, and tall forbs
		F3	Temperate & Mediterraneo-montane scrub habitats
		F9	Riverine & fen scrub
		G1	Broadleaved deciduous woodland
		G2	Broadleaved evergreen woodland
		G3	Coniferous woodland
		G4	Mixed deciduous and coniferous woodland
		G5	Lines of trees, small anthropogenic woodlands, recently felled woodland, early-stage woodland and coppice
2	Coniferous woodland	G3	Coniferous woodland
		G5	Lines of trees, small anthropogenic woodlands, recently felled woodland, early-stage woodland and coppice
3	Boundary and linear features	FA	Hedgerows
		G5	Lines of trees, small anthropogenic woodlands, recently felled woodland, early-stage woodland and coppice
		H5	Misc inland habitats with very sparse or no vegetation
		J2	Low density buildings
		J4	Transport networks and other constructed hard-surfaced areas
4	Arable and horticultural	E2	Mesic grasslands
		I1	Arable land and market gardens
5	Improved grassland	E1	Dry grassland
6	Neutral grassland	D6	Inland saltmarshes
		E2	Mesic grasslands
		E3	Seasonally wet & wet grasslands
7	Calcareous grassland	E1	Dry grassland
		F2	Arctic, alpine & subalpine scrub habitats
8	Acid grassland	E1	Dry grassland
		E3	Seasonally wet & wet grasslands
9	Bracken	E5	Woodland fringes & clearings, and tall forbs
10	Dwarf shrub heath	F4	Temperate shrub heathland

11	Fen, marsh and swamp	C3 D2 D4 D5 E5	Littoral zone of inland surface waterbodies Valley mires, poor fens and transition mires Base-rich fens and calcareous spring mires Sedge and reedbeds, normally without free-standing water Woodland fringes & clearings, and tall forbs
12	Bogs	D1	Raised and blanket bogs
13	Standing open water and canals	C1 C3 J5	Surface standing waters Littoral zone of inland surface waterbodies Highly artificial man-made waters and associated structures
14	Rivers and streams	C2 J5	Surface running waters Highly artificial man-made waters and associated structures
15	Montane habitats	E4 F2	Alpine & Sub-alpine grasslands Arctic, alpine & subalpine scrub habitats
16	Inland rock	E1 E5 H1 H2 H3 H5 J3 J6	Dry grassland Woodland fringes & clearings, and tall forbs Terrestrial underground caves, cave systems, passages and waterbodies Screes Inland cliffs, rock pavements and outcrops Misc inland habitats with very sparse or no vegetation Extractive industrial sites Waste deposits
17	Built up areas and gardens	I2 J1 J2 J4 X22 X23 X24 X25	Cultivated areas of gardens and parks Buildings of cities, towns and villages Low density buildings Transport networks and other constructed hard-surfaced areas Small city centre non-domestic gardens Large non-domestic gardens Domestic gardens of city and town centres Domestic gardens of villages and urban peripheries
18	Supralittoral rock	B3	Rock cliffs, ledges and shores, including the supralittoral
19	Supralittoral sediment	B1 B2	Coastal dunes and sandy shores Coastal shingle
20	Littoral rock	A1	Littoral rock and other hard substrata
21	Littoral sediment	A2	Littoral sediment

Annex 3

UK Plant Surveillance Workshop: agenda and notes, February 2010

UK Plant Surveillance Workshop
Board Room, Plantlife International, Salisbury SP1 1DX
Thursday 18th February 2010
9.30am - 4.30pm

Agenda

- Welcome and introductions (Trevor)

- Frame of reference (Nicola)
  Summary of discussion group
  JNCC terms of reference
  Overview of current schemes
  Volunteer engagement

- Stratification of 1km squares including overlay with other schemes (Kevin)

- Survey methods (Kevin)
  Distribution, abundance, sampling within 1km squares (plots v transects)

- Habitat stratification (Trevor)
  Broad Habitat, NVC, Phase 1, EUNIS

- Selection of target species (Trevor)
  Frequency, habitats, traits, ease of identification

- Optimal number and time of visits (Kevin)

- Parameters to record (Trevor)
  Vegetation structure
  Other taxonomic groups

- Summary discussion and next steps (Nicola)

Attendees:

Alison Johnson, BTO; Chris Preston, CEH; David Noble, BTO; David Roy, CEH; Kevin Walker, BSBI; Nicola Hutchinson, Plantlife; Stephen Freeman, CEH; Tim Pankhurst, Plantlife; Trevor Dines, Plantlife

Introduction

The workshop discussion was based around a series of prepared papers. Below are a series of brief notes on the main questions posed and discussions held. Much of the discussion focused on the practicalities of implementing the methods being reviewed, the conclusion being that field testing is essential to aid the selection of a single surveillance method.

Stratification of 1km squares including overlay with other schemes

1. *Are 1km squares suitable?*

One kilometer squares are useful for selecting the sample – other schemes and datasets use 1km squares - but should this be the actual unit of survey? Potential use of transects and plots (nested or otherwise) within the 1km square.

Discussion on concerns around need to record all species in entire square to truly pick up environmental changes occurring. Vegetation recording at ever smaller scales results in a greater number of habitat/vegetation dominants being recorded, so ideally the greater area surveyed provides a better representation of species within and across vegetation/habitat categories. However, overall advice that do not need to record every species across an entire 1km square in order to pick up change. Sample size and repeatability will allow trends to emerge, so long as the survey is seeking to record something which is considered to be proportionally representative of wider environmental changes.

Butterfly scheme – specifically added common species survey to detect wider countryside changes because more targeted specialist species recording was not delivering this information.

2. *Are VCs adequate geographic strata or would environmentally defined strata be more appropriate?*

Botanical recording has in the past been stratified to give full geographic coverage, whereas other taxonomic recording schemes (e.g. Breeding Bird Survey) stratify according to potential recorder density within an area.

Can take a random sample of squares to begin with but experience of butterfly scheme is that recorder choice will bias the sample since some squares are more popular to record than others.

Use general population density (to weight samples) and sample accordingly within any given geographic area. Start with a lower sample size in high population density areas than ultimately required, and build up participation over time.

Geographic strata are impossible to avoid so might as well use, rather than trying to use landscape or other environmental unit

Pros of using the Vice County network include = static network; promotes ownership of squares for BSBI recorders. Cons include = some are quite small areas; issue of smallest Vice County with greatest population (Middlesex)

Consider issue of the distances recorders would travel to survey squares in areas which need gap-filing. Consider potential recorder base to include BSBI regular surveyors, other BSBI members, Plantlife members, other Plantlife survey participants, other scheme participants (e.g. Breeding Birds, Butterfly Scheme).

3. *Is it sensible to use existing squares?*

4. *Can non-random squares be included in the 'pool' of random squares (e.g. Local Change)?*

Can use different sets of squares to make up the total sample and perform statistical analysis to check for any differences in results between ‘schemes’. The butterfly monitoring scheme uses a randomly generated sample for ‘original’ butterfly scheme recorders, plus the Breeding Bird Survey squares.

Local Change squares could be considered randomly stratified so long as the origin of the sample grid is considered to be random (i.e. the OS grid). Issue with Local Change squares would be the size (tetrad), although a system could be applied (e.g. the SW 1km square within each tetrad).

Statistical checks needed to ensure that sample sizes are adequate from any scheme.

Consider how to ensure spread of recorder effort with ‘pooled’ samples – allocation of squares to surveyors.

5. *The recorder weighting biases samples to more populous regions - is this a problem?*

Covered in discussion under question 2.

Expect density of organizational membership to reflect population density – doesn’t matter ultimately since will vary from that (BBS example) due to behavior of volunteers (BBS filled dense areas more quickly and needed to allocate more within those areas, yet some areas never recorded).

Question of how to aggregate (need to wait year or two) – so might have a bias in the first year but build up to fill gaps over years. Start with low allocation in high density areas and build up (start with a smaller sample and add other random samples in time).

6. *When weighting what do we mean by 'recorders' (active, members, total population density)?*

Again, covered under discussion under question 2.

Need to understand the minimum number of 1km squares required per Vice County (or other chosen geographic unit) based on total population density. Accept that will get more survey effort in denser population areas; areas with lower population density/remote access may need to be undertaken by professionals or particularly dedicated volunteers - can consider concept of ‘super squares’ (i.e. adjacent squares) to save time traveling.

Survey methods

1. *What is our preferred method?*
2. *Is it more important to be representative or target habitats regardless of extent in square?*
3. *If using transects is bias to linear features a problem?*
4. *Could transects be linked to other schemes?*
5. *Will volunteers record abundance in small, fixed plots?*
6. *What 'indices' of change could be derived from these methods?*

The workshop discussed the details of methods proposed and voted. Methods 5 and 7 (latter = new method proposed by Simon Smart) were most popular, although further discussion on method 5 suggested less differences than perceived between this and method 3. Methods 5 and 3, and method 7 considered those to take forward into the statistical analysis stage and for further consideration by the project team.

Specifics from the discussion:
Method 2 – idea of using random lines of grid cells (horizontal or vertical) instead of random scatter of grid cells. Recorders will need to deviate from the line due to access issues in places; also possible issue of habitat clustering.

Method 4: issue with 2m being little more than the width of the path. Will interesting species outside of the survey area be recorded - is the data used? Take lessons from other schemes.

Method 5 – forces people off the footpaths. Could record all species along path and more detail in the plots. Could you randomly select cells along transect rather than within grid cell – because there is the issue of ownership (paths = route of least resistance). Concern was expressed that the statistical bias from using paths is greater than the risk of losing/putting off participants, although the point was made that you can get bias in other ways. BBS is mostly done from rights of way and when looking at broad habitat classes don't detect a significant effect. Also, discussion on other surveys (including Countryside Survey) often undertaken using rights of way yet data still statistically valid. Method 5 gets around the path problem since you use rights of way to get around 1km sq but plots are then less biased to rights of way.

Do not need to measure species abundance for the aims of the project, frequency of occurrence is enough. Just need relative abundance (not absolute abundance) to measure status and trends. So don't have to use quadrats that would measure absolute abundance (CS doesn't use the abundance measure). With frequency we mean how often a species occurs i.e. measuring relative freq within a 1km^2.

Method 7 (SS): method ADAS trialed when monitoring ESAs - nested plot approach applied at 1km^2. Come up with species list for the 1km ^{2} using an efficient search regime based on starting with a small plot in corner of square and then only record additional species in increasing sized squares. At each occurrence of new species also tick box for broad habitat so have species and habitat cumulative curves (reflects habitat diversity within square by habitat curve plus species occurrence curve). Will not give more than presence/absence in 1km^2 but will indicate where species are more or less common where there are present across UK. Method relies on being able to move freely around square i.e. ignores the ownership issue. How to analyse the data – could map numbers of species in each nest, plus species list for whole square. Would work for a subset of species, but when recording all species brings in rarer species which may have key signals of change.

Does the survey need to be completed in one day, or could surveyors spread the survey over more than one day? Concern that this might put some off. Could do different squares in different years to keep interest.

Stephen Freeman introduced a time-species count of birds method, and suggested a spatial equivalent for use in this project. Bird method was used in Uganda because lots of species and few people. Stephen presented a paper and agreed to follow up with further explanation.

Issues with methods 3/5 to overcome:

- Nobody is actually used to surveying a 100mgrid cell – rather used to taking 6fig GR of a species – need to explain new method.
- Do volunteers need to record dull squares comprehensively? Or actually ‘get done’ with dull squares and move on – is this robust?
- Which method is easiest to implement? Need field testing
- Consider the use of a small pool of volunteers who would record all day or larger number of volunteers who will only spend a few hours, or mix of both.
- Which statistically better at detecting change (regardless of volunteer effort)?
- Access issue: use of linear features (-ve), trespass or gain access?
- Use comparison with Countryside Survey.
- Must link species observation to habitat.
- Plots are resource intensive e.g. volunteer support & relocation issues.
- Minimized support costs with method 7 since surveyors are free to choose route. Could do along a walk but path issues etc to take into account.

Field testing: trial different methods in the same square; restrict recording to rights of way; compare Countryside Survey to the favoured methods.

Understanding that implementation of schemes is key to determining which method is preferred. Implementation includes potential for volunteers to be enthused and committed, and level of organisational support required

Habitat stratification

1. *Do we adopt Broad Habitats as the basis for habitat stratification?*

2. *If so, do we exclude BHs with few species (e.g. Bracken) or with few widespread species (e.g. montane)?*

3. *Do we need to sub-divide some BH categories, perhaps using Eunis divisions as above? If so, how do we reach consensus over sub-divisions and subsequent species classification?*

Agreement on the use of EUNIS as the habitat classification scheme, rather than Broad Habitats, although some more 'friendly' names will be required to translate for volunteers.
Discussion on the possibility of BRC mapping species to EUNIS habitats
JNCC have produced a spreadsheet to standardise Broad Habitats against EUNIS.

Selecting target species

Recording all species is better technique to show trends, however probably not possible as an annual survey. Plants don't show rapid change unless look at fluctuating annuals. But trends do build up over years.

Two tier survey acceptable e.g. annual recording of a target list of species and less regular recording of all species.

If looking for big signals of change over long periods of time – need only perhaps 200-300 squares with all species recorded, and 1000 with target list. Local change suggests 750hectads is a useful cut off point for having covered most species.

In terms of rotation (i.e. all species recorded once every three or five years) there will need to be a repeat before trends come through so results only possible after 6 or 10years. Alternative would be to record all species in five squares each survey on a five year rotation, but again need to wait ten years before first results emerge.

1. *How many levels of target species-list do we have? Three are proposed, but would two or four be better?*

Two: easy and expert.

Those who can confidently record 500 species will record those less well than those that do all since will have lower field skills.

Consider observer effect – i.e. learning curve.

2. *Should we select species that are representative of a range of broad habitats or drawn from species that are indicative of particular broad habitats?*

Choose species easier to identify and indicative of habitat, although prefer a mix of specialist and generalist species plus any specifically relevant for other taxonomic groups. In terms of policy influence from the butterfly scheme, specialists show a decline compared to generalists, so is useful to have recorded both types. Check whether any ecosystem functions are missing from species list to complete range.

Optimal visits

Can the average flowering time be used in the analysis to know whether positive or negative changes are seasonal changes – if you can correlate trends against average flowering time.

1. *Are we assuming a single visit or could survey work be spread over the season?*

2. *Recorders will not know habitats a priori. Therefore, would we need to categorise 1kms by habitats prior to survey (using species distribution data)?*

3. *One kilometer would contain more than a single BH - therefore, how do we cope with more than one peak month? Just say all survey should take place June-August (BL wood, saltmarsh?)*

4. *Should we be concerned with regional variation?*

5. *Does the data from each visit need to be treated separately?*

Two visits in lowlands (to pick up woodland in Apr/May), peak visits in June – also consider winter annuals. Depends on the type of volunteer whether they need to have species in flower.

Do not need to categories 1km² by habitat in terms of visit, but need to think about in terms of whether the 750 hectad cut off point is changed by country size. Do you need more samples within country to make national trends relevant (e.g. CS in Wales).

Does data from separate visits need to be treated separately – suggestion that we only want a single form/set of data (not a problem if abundance being excluded). Two visits for BBS – just use the one that's most appropriate for the species. Ask them to say how long spent on site surveying – can vary over time.

Parameters to record

1. *How do we record percentage cover and/or frequency in any of the suggested methods?*

In defence of % cover bands – can turn category into a % and don't need too much accuracy and it can be collapsed down to presence/absence. Presence/absence can be combined with % cover – think it would be a shame to sacrifice % cover bands. But cannot put % cover into large plots.

With need to measure small scale changes concerned about only species presence/absence at one square kilometre.

DAFOR generally agreed as useful with some measure of local abundance factored in.

2. *Is this an optimal set of attributes to record?*

Broad agreement with suggestions made, although field testing required to see how easy they are for volunteers to record. Keep simpler if longer list is a barrier to participation or retention.