Gaining new knowledge from historic data: an approach to ecological data rescue, with special reference to UK Centre for Ecology & Hydrology (UKCEH) long-term land use monitoring data sets

Claire M. Wood

PhD (by Research Publications)

The University of Edinburgh

2020

Declaration

I declare that the thesis presented for the degree of PhD (by Research Publications), has been composed by myself and that the work has not been submitted for any other degree or professional qualification. I confirm that the work submitted is my own, except where work which has formed part of jointly-authored publications has been included. My contribution and those of the other authors to this work have been explicitly indicated in Section 4. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

Parts of this work have been published in Earth Systems Science Data and the Journal of Landscape Ecology.

Claire Wood

Acknowledgments

Firstly, I would like to thank Bob Bunce, for ongoing encouragement, support and advice over the last few years, and for inventing the mad world of the Countryside Survey in the first place. Thanks Bob – I owe you a beer!

I would also like to thank the following:

Marc Metzger for valuable guidance and support throughout the production of this document this year.

Members of the Land Use Group past and present, in particular Simon Smart, Peter Henrys, David Howard, Lisa Norton, John Watkins and Rick Stuart.

Annali Bamber-Jones and Sue Ward for support, encouragement and advice.

Fellow ILs.

The UK Centre for Ecology & Hydrology for supporting this work and for allowing use of photos from the archive.

Abstract

The overarching theme of my collection of work comprises a unique synthesis of environmental informatics and long-term ecological monitoring, set in a context of a period of rapid developments in field survey methodology and environmental change. It concerns ecological data from a regional to national scale, charting the development of the repeatable long-term ecological monitoring undertaken by the UK Centre for Ecology & Hydrology (UKCEH) (formerly the Institute of Terrestrial Ecology, ITE). My submission consists of six peer-reviewed first author papers (plus an additional second author paper) representing a major exercise, led and largely undertaken by myself, in data rescue, management, analysis and publication of a series of nationally important ecological data sets. Each of my published papers explores a major ecological survey, describes the methodology, available data and findings, and places each survey into a national, and international, context.

My collection of work demonstrates my technical expertise in the sphere of ecological data management and has been pioneering in the field of data publication and open data. There is an increasing movement, encouraged by the UK Research Councils, towards making publicly funded research data openly accessible in a data repository. In recent years, several journals have been set up to encourage the sharing and re-use of scientific data. Publishing my work in such journals has meant that nationally valuable data are openly available and described to the wider research community and the wider public ensuring data transferability and re-use. Without my vital work concerning the environmental data described, the quality assured data would not be secured safely for long-term storage and use. Indeed, some of the data sets were already on the verge of being lost before the work in question was undertaken.

My papers cover a period beginning in the late 1960s and early 1970s, before which sampling techniques for ecological survey tended to be subjective and non-repeatable. Thanks to the ground-breaking work undertaken at the Institute of Terrestrial Ecology's research station at Merlewood, Cumbria, statistically robust, standardised, repeatable sampling methods have been developed for producing figures for large areas. My papers explore how these methodological techniques began to evolve in the first two surveys examined, concerning two habitat specific woodland surveys undertaken in 1971 (across the whole of Great Britain and in Scottish Pinewoods). These were followed by a regional survey of Shetland (1974), during which the idea of a statistical sampling framework using stratified random sampling was first tested. This led to the initiation of the first national ecological survey of Great Britain in 1978, now arguably the largest long-term ecological monitoring project in the country, known as the Countryside Survey (CS). This national survey has now been repeated in 1984, 1990, 1998 and 2007 and is covered in the final three papers. An additional paper describes how the techniques were utilised in an associated survey of targeted 'Key Habitats' in England in the 1990s, one of several surveys associated with Countryside Survey.

The countryside of Great Britain and its associated habitats and ecological features have changed considerably over the last 50 years, for a variety of reasons. My collection of work provides the unique opportunity to explore the changes and drivers of change that have taken place in the British countryside. The published data enable links between different disciplines to be made, furthering a range of research on many aspects of land use and land use change

such as evolving farming and forestry practices, climate change and atmospheric pollution and providing evidence for national policy makers. The final paper included in my submission is a specific example of how this type of data may be applied, in this case in relation to the British uplands.

This document should form essential reading for all scientists planning a major ecological project involving repeatable measurements.

Lay summary

My collection of work comprises a unique combination of environmental informatics and long-term ecological monitoring, set in a context of a period of rapid developments in field survey methodology and environmental change. It concerns long-term ecological monitoring data collected by UK Centre for Ecology & Hydrology (UKCEH) throughout Great Britain since the 1970s, spanning a time of considerable change in the countryside and its associated habitats and ecological features.

My submission consists of six peer-reviewed first author papers (plus an additional second author paper) representing a major exercise in data rescue, management, analysis and publication of a series of nationally important ecological data sets, on the verge of being lost before this work.

The work has been pioneering in the field of data publication and open data. Publishing my work in open access, data-oriented journals has meant that nationally valuable data are openly available and described to the wider research community and the wider public ensuring data transferability and re-use. Without my vital work, these data sets could have been lost forever. My collection of work facilitates the unique opportunity to explore the changes and drivers of change that have taken place in the British countryside over time.

This document should form essential reading for all scientists planning a major ecological project involving repeatable measurements.

Table of Contents

Critical Review

Critical Review

1. Introduction and research context

1.1 Overview

This critical review explores a body of work related to a unique series of nationally, and internationally, significant long-term (or potentially long-term) ecological monitoring data sets concerning land use and land cover, originating from the UK Centre for Ecology & Hydrology (UKCEH) and its predecessor bodies, arising from field surveys initiated by Professor R.G.H. Bunce from the 1970s onwards. My six peer-reviewed first author papers (plus an additional second author paper) represent the culmination of a major exercise in data rescue, led and largely undertaken by myself. The rescue tasks involved a range of activities related to data management, documentation, publication of data products and data analysis. Each of my published papers explores a major ecological survey (or an element thereof), describes the methodology, available data, metadata and findings, and places each survey into a national, and international, context.

The first part of this review explores the significance of what has become termed the 'Bunce' ecological surveys and data. The review is then set into the context of the ascent of 'open data' and 'open access' initiatives. There is an increasing movement, encouraged by the UK Research Councils, towards making publicly funded research data openly accessible in a data repository (Collins, 2011; Molloy, 2011; UKRI, 2020). The 'Bunce' UKCEH data sets were among some of the first to be deposited in the Natural Environment Research Council (NERC) Environmental Information Data Centre (EIDC), contributing to the development of procedures used for doing so. In recent years, several journals have been set up to encourage the sharing and re-use of scientific data. Publishing my work in such journals has meant that nationally, and internationally, valuable data, and importantly, metadata, is openly available to the wider research community and the general public ensuring data transferability and re-use. Without my vital work concerning the environmental data described, the quality assured data would not be secured safely for long-term storage and use, indeed some of the data sets were already on the verge of being lost before the work in question was undertaken. These 'Bunce' data sets constitute the most comprehensive large-scale ecological time series in existence. An example of how this type of data may be used is demonstrated in Paper 7 (see Section 4 for list of numbered papers).

The originality of my work lies in the approach to the data set rescue, reconstruction, documentation and publication, thereby transforming collected field data into a re-usable product. Data are the most important output from scientific studies, and yet managing the data properly is the most overlooked aspect (Molloy, 2011; Whitlock, 2011). Data sets should be considered as first class research outputs, comparable to peer-reviewed journal articles (Callaghan et al., 2012; Costello et al., 2013; Lawrence et al., 2011). It has been recognised that 'the distance between data collectors and those skilled in delivering the data is considerable', and this is further accentuated when data delivery is temporally distant from data collection (Specht et al., 2018), with the original survey data at risk of being lost. My work has entailed tackling the technical challenges and primary source research, which are described in detail. My collection of work demonstrates my technical expertise in the sphere of ecological data management and has been pioneering in the growing field of data publication and open data. Although much progress has been made in the past 20 years towards the routine publication of data, properly described, protected and archived for future use, the recovery of past ecological data remains in its infancy (Costello et al., 2013; Downs & Chen, 2017; Specht et al., 2018).

Finally, conclusions are drawn as to the wider context and significance of the data rescue exercise and approach, as well as the lessons learnt for future data collection and the legacy of the data associated with the surveys.

1.2 Background to UKCEH surveys

UKCEH is fortunate in having a relatively long history of ecological recording, leading to a wealth of historic resources. UKCEH partly traces its origins back to the Nature Conservancy, set up by Royal Charter in 1947 (with its first research station being opened in 1952 at Merlewood in Cumbria). This partly evolved into the Institute of Terrestrial Ecology (ITE) in 1973 (as a component of NERC), then the NERC Centre for Ecology & Hydrology in 2000, finally transitioning to an independent body with charitable status in 2019 as UKCEH, with four research stations. Within this timespan, with research stations across Britain, UKCEH has been at the forefront of ecological research in the UK and further afield, and has been uniquely placed to carry out national long-term research and monitoring. In addition to research on UK land cover and land use, well-known projects and data collection activities include the Predatory Bird Monitoring Scheme (PBMS), the Biological Records Centre (BRC) and the Environmental Change Network (ECN), among many others (UKCEH, 2020b).

One key area of strength has been research into land use and land cover across the UK. Originating in the 1950s as a small woodland research group based at the former Merlewood Research Station, over time this developed into the Land Use Research Group (now based at UKCEH Lancaster), studying land use of all kinds by the late 1970s. It is now perhaps most well-known for managing the unique long-term study known as the Countryside Survey and the UK Land Cover Maps (LCM) which accompanied Countryside Survey until 2007, and are now produced on independent timescales (UKCEH, 2020c). My papers cover a period beginning in the late 1960s and early 1970s, at a point when most projects involving ecological survey tended to be subjective and non-repeatable. Thanks to the pioneering work undertaken at Merlewood, statistically robust, standardised, repeatable sampling methods were developed for producing estimates of ecological metrics for large areas, developing into the national ecological survey, the Countryside Survey (Sheail & Bunce, 2003).

1.3 The 'Bunce' series of data sets

The 'Bunce' series of data sets all follow a similar methodological approach and design, as introduced in the 1970s. Early surveys in British woodlands (Papers 1 and 2) and Shetland (Paper 3), and also Cumbria (Bunce & Smith, 1978; Bunce et al., 2017c), led to the initiation of the first national ecological survey of Great Britain in 1978, now arguably the largest and most comprehensive long-term ecological monitoring project in the world, the Countryside Survey. This national survey has been repeated successfully in 1984, 1990, 1998 and 2007 and is covered in Papers 4, 5 and 7. The latest repeat began in 2019 as a 'rolling' survey over a five year period, focusing on soils and vegetation. An additional paper (6) describes how the techniques were utilised in an associated survey of targeted 'Key Habitats' in England in the 1990s, one of several surveys associated with Countryside Survey.

The countryside of Great Britain and its associated habitats and ecological features have changed considerably over the last 50 years, for a variety of reasons. The work I have undertaken facilitates a unique opportunity to explore the changes and responses to drivers of change that have taken place in the British countryside. The published data enable links to be made between different disciplines, furthering a range of research on many aspects of land use and land use change such as evolving farming and forestry practices, climate change and atmospheric pollution and providing evidence for national policy makers. The final paper (Paper 7) included in my submission is a specific example of how this type of data may be interrogated, in this case providing a characterisation of vegetation in the British uplands.

1.4 Stratified sampling strategy and 'Bunce' survey design

As described by Sheail and Bunce (2003), there have been many attempts to develop ecological classificatory systems, ever since early vegetation maps at the turn of the twentieth century. In Britain, early approaches focused on dominant species and the general state of plant communities (Moss, 1910; Tansley, 1939). Later developments classified vegetation using characteristic species (Braun-Blanquet, 1932; Poore, 1955). These approaches relied upon subjectivity and the intuition of surveyors, and it became increasingly clear among ecologists in the 1950s that their science needed a more exact basis (Webb, 1954), moving towards standardised systems that would enable an evaluation of change over time or differences across space that was independent of opinion (Sheail, 1987; Sheail & Bunce, 2003).

At Merlewood, the woodland research team recognised a need for a National Woodland Classification, providing an obvious opportunity to work on a quantitative, statistical approach to vegetation survey in the Nature Conservancy (Sheail & Bunce, 2003). Thus, principles of strategic classification and subsequent sampling from defined strata were initially devised for broadleaved woodland, and first tested at a regional level in Shetland in 1974 (Wood & Bunce, 2016b). To a greater or lesser extent, all of the 'Bunce' surveys use this stratified sampling strategy, based on the same statistical principles. They rely on classical regression theory, with the environmental classification being the independent, and the vegetation or habitat the dependent, variable (Sheail & Bunce, 2003). In brief, the approach consists of an environmental classification stage which involves the creation of a stratification (or areas of relatively homogenous regions) of the area in question (for example, Cumbria, Shetland or Great Britain) using multivariate classification of environmental characteristics (based on the principles outlined by Hill et al. (1975); Williams and Lambert (1959)). In order to sample ecological parameters such as vegetation, sampling sites are then chosen at random from within the strata created. By using this statistically robust method, it is then possible to make statistical estimates from the sample sites to describe the entire population with associated error terms. By the end of the 1970s, this idea of stratifying the landscape led to the successful creation of a stratification for the whole of Great Britain, known as the 'Institute of Terrestrial Ecology (ITE) Land Classification of Great Britain' (Bunce et al., 1990; Bunce et al., 1996a). Although this has developed over time (Bunce et al., 1998; Bunce et al., 2007), the basic stratification still underpins the Countryside Survey as described by Carey et al. (2008) and Norton et al. (2012b), and the statistical principles are described in detail by Metzger et al. (2013) as well as its possible extension to Europe.

By 1990, the Countryside Survey integrated programme of field survey and remote sensing had been developed in the UK, which had the potential to provide a statistical summary of changes in the land cover and countryside features of the whole country every ten years (Barr et al., 1993). An outline of each of the UKCEH 'Bunce' surveys in question is given below.

1.4.1 Woodland Survey of Great Britain (Paper 1)

The Woodland Survey of Great Britain is a unique data set, consisting of a detailed range of ecological measurements at a national scale, covering a time span of 30 years. A set of 103 woodlands spread across Great Britain were first surveyed in 1971 by the Nature Conservancy, which were again surveyed between 2000 and 2003. Standardised methods of describing the trees, shrubs, ground flora, soils and general habitats present were used for both sets of surveys. The sample of 1648 plots spread through 103 woodland sites located across Britain makes it the most extensive quantitative ecological woodland survey of its type undertaken in Britain; it is also notable for the range of sites that have been revisited after such a long interval. The 103 surveyed woodlands were selected from a set of 2453 woodlands that had been part of a preliminary survey known as the 'Steele' survey (Steele, 1968). This had begun in the late 1960s and was led by R. C. Steele, the head of the Nature Conservancy's Woodland Management section. Standard recording cards were used, and the data provided background information for the Nature Conservation Review (Ratcliffe, 1977). The subset of 103 was derived from the 2453 by Association Analysis (Williams & Lambert, 1959) and other numerical techniques that, at the time, were still novel and undergoing rapid development (Bunce et al., 1981; Bunce & Shaw, 1973; Hill et al., 1975). These analyses placed the woodlands into 103 groups according to the similarity of their plant species composition. The woodland that was most typical of that group was then selected for detailed survey, using Principal Components Analysis. Later analysis shows the sites chosen for survey are proportionally representative of the woodland area in each strata in the ITE Land Classification. Before the rescue work, these data were not publicly available.

1.4.2 Ecological survey of the native pinewoods of Scotland (Paper 2)

In 1971 (the same year as the Woodland Survey of Great Britain), a comprehensive ecological survey of the native pinewoods of Scotland was also carried out. The survey was initiated as a consequence of growing concern about the status of the pinewood resource. Since the twentieth century, this unique habitat has been widely recognised, not only by ecologists for its inherent biodiversity but also by the general public for its cultural and amenity value.

This survey varies slightly from the other 'Bunce' surveys in terms of design, in that nearly the entire population of the habitat was sampled. The population sampled was the major 27 sites of the 35 sites identified as truly native pinewoods in Scotland by Steven and Carlisle (1959). However, the project utilised the same repeatable methods as used in the Woodland Survey, collecting information on ground flora, soils, forest structure and also general site information for each site. The results from the survey prompted the organisation of an international symposium in 1975, which set the conservation agenda for the old Caledonian pinewoods (Bunce & Jeffers, 1977). Although the data are now 49 years old, the repeatable methods allow a resurvey to take place, beginning in 2020, in order to assess changes in the vegetation, habitats and tree composition in a statistically robust manner. Before the rescue work described, these data were not publicly available or in a usable state.

1.4.3 Survey of the terrestrial habitats and vegetation of Shetland (Paper 3)

A survey of the natural environment was undertaken in Shetland in 1974 by ITE (Milner, 1975), commissioned after concern was expressed that large-scale development from the new oil industry could threaten the natural features of the islands. A framework was constructed by ITE on which to select (1km square) samples for the survey, then vegetation and habitat data were collected. In addition to providing valuable information about the state of the natural environment of Shetland, the repeatable and statistically robust methods developed in the survey were used to underpin the national Countryside Survey, being for the entire area of land, rather than solely woodland habitats, as previously in the two woodland surveys. Whilst a repeat has not yet taken place, the demonstration of the effectiveness of the methodology in other surveys indicates that a repeat of the Shetland survey would yield statistics about ecological changes in the islands, such as those arising from the impacts of the oil industry, a range of socio-economic impacts, and perhaps climate change. Currently no such figures are available, although there is much information on the sociological impacts, as well as changes in agriculture. Before the rescue work, these data were not publicly available or in a usable state.

1.4.4 Countryside Survey of Great Britain (Papers 4,5,7)

As already described, Countryside Survey is the biggest and most repeated of the 'Bunce' surveys. Countryside Survey is based on 1 km squares as a convenient sized unit, which had previously been tested in Cumbria (Bunce & Smith, 1978) and Shetland (Wood & Bunce, 2016b) in the years preceding the first national survey in 1978. The survey design is based on a series of dispersed, stratified, randomly selected 1 km squares from across Britain, which numbered 256 in 1978, 506 in 1990, 569 in 1998 and 591 in 2007. The stratification used was the statistical environmental classification of 1 km squares in Great Britain as described in Bunce et al. (1996a, 1996b).

In addition to vegetation species from plots, data were also recorded from linear features such as hedgerows, landscape elements such as veteran trees, areal broad habitats (Jackson, 2000) and related key species, soils and freshwaters (see Carey et al., 2008). The survey as a whole provides a wealth of globally unique ecological data, consisting of an extensive range of measurements at a national scale, covering a time span of 29 years. From an international perspective, Countryside Survey was a pioneer in surveys of its type. The integrated, systematic national monitoring of vegetation species, soils and landscape features across all land uses provided by Countryside Survey was a novel concept, preceding programmes in many other countries particularly in Europe (for example Dramstad et al. (2002); Hintermann et al. (2002); Ståhl et al. (2011) and beyond (for example Burton et al. (2014)). Currently a partial repeat survey is in progress, having begun in 2019 as part of a five-year rolling programme.

Whilst most of the data collected from Countryside Survey have not been at risk at any point (being well managed as one of UKCEH's key data sets), the emphasis of the work required was on documenting and publishing the data, and on data rescue for data collected in 1978 survey which were not publicly available or in a re-usable state.

1.4.5 Ecological survey of `Key Habitat' landscapes in England (Paper 6)

The sampling framework for the national Countryside Survey is not optimised to yield data on rare or more localised habitats. In the 1990s, a survey was commissioned by the former Department of the Environment (DoE, now the Department for Environment, Food and Rural Affairs, DEFRA) to carry out additional survey work in English landscapes which contained semi-natural habitats that were perceived to be under threat, or which represented areas of concern to the ministry. The landscapes were lowland heath, chalk and limestone (calcareous) grasslands, coasts and uplands. The information recorded allowed an assessment of the extent and quality of a range of habitats defined during the project, which can now be translated into standard UK broad and priority habitat classes (Jackson, 2000; Maddock, 2008). The survey, known as the 'Key Habitat Survey', followed a design which was a series of gridded, stratified, randomly selected 1 km squares taken as representative of each of the four landscape types in England, determined from statistical land classification and geological data ('spatial masks'). A total of 213 of the 1 km square sample sites were surveyed in the summers of 1992 and 1993, with information being collected on vegetation species, land cover, landscape features and land use, applying standardised repeatable methods. The database has contributed additional information and value to the long-term monitoring data gathered by Countryside Survey and provides a valuable baseline against which future ecological changes may be compared, offering the potential for a repeat survey. Before the rescue work, these data were not publicly available or in a usable state.

1.5 Originality of the work

Whilst it is clear that the original 'Bunce' surveys are unique and pioneering in many ways in terms of survey design and implementation, it is essential to emphasise how the current body of work in question is original in itself. Gill and Dolan (2015) identify several areas of identifying originality in the context of a PhD. In this case, two of these areas in particular apply to this body of work. Firstly, the work provides 'a synthesis of information that has never been put together before' and secondly, the work 'adds to knowledge in a way that has not been done before'.

1.5.1 Novel synthesis of existing information

An important point to make is that there is a recognition that 'the distance between data collectors and those skilled in delivering the data is considerable' (Specht et al., 2018).

Overall, the work relates to ecological informatics, which is an interdisciplinary field that includes conceptual and methodological tools for the understanding, generation, processing and dissemination of various types of ecological data (Michener et al., 2002). Within a project or study, ecological informatics contributes to: (I) Experimental design phase; (II) Data design plan; (III) Data acquisition and management; (IV) Quality assurance and control (QA/QC); (V) Metadata implementation; (VI) Data archival; (VII) Data access and dissemination; (VIII) Data publication and (IX) Analysis. This framework has been described as the 'data life-cycle approach' (Alves et al., 2018; Michener & Jones, 2012).

Table 1.1. My contribution to the stages of the 'Bunce' UKCEH research projects (as proposed by Michener and Jones (2012)) in terms of ecological informatics

Stage of a research project	***Original scientists' contribution***	***My contribution***
(I) Experimental design phase	*	
(II) Data design plan (database design)		*
(III) Data acquisition and management (in field, and later rescue)	*	*
(IV) Quality assurance and control (QA/QC)	*	*
(V) Metadata implementation		*
(VI) Data archival		*
(VII) Data access and dissemination		*
(VIII) Data publication		*
(IX) Facilitation of data analysis / undertaking analysis	*	*

Of these nine stages, my work described here encapsulates the majority of stages II to IX in relation to the UKCEH 'Bunce' surveys (Table 1.1) to a greater or lesser degree depending on the specific data set in question. In this case, Stage II is specifically referring to design of the data in a relational database and Stage III includes both the acquisition of the original data in the field (original scientists) but also the re-acquisition of data at the point of rescue (my contribution). The data analysis phase includes working towards the provision of well-managed data to facilitate analyses personally and for other scientists and data users (also including the analyses undertaken at the time of the surveys by the original scientists). Further detail on my specific contribution to the work relating to each publication is outlined in Sections 3 and 4.

Although the original surveys were designed with longevity and repeatability in mind, the technological tools and skills to deal with the practicalities of this in terms of data and metadata were rudimentary or non-existent at the time, leaving the key data and information required for the undertaking of a repeat survey at the risk of being lost forever. Although the original survey documents (including data recording sheets, field handbooks, code sheets, notes, maps and photographs) were fortunately put in storage at the time of the surveys, little consideration was given as to how the information and data could be of practical use in the future; staff had moved on to other projects.

Producing published data products (as a discrete data and metadata package), as is becoming commonplace today, was beyond the capabilities of contemporary technology for these surveys. The data and metadata and also the methods from these surveys had not previously been an output from the work; much of the early data were collected on paper and then never digitised or electronically stored in a way that endured. Preparing the data for re-use has not been a simple matter; it has been a complex task of data reconstruction. A considerable amount of work gathering information from primary sources such as paper archives, digital archives and interviews with personnel from the original work has been carried out by myself, as well as a great deal of data manipulation and management. The emphasis is on the creation of the data products and metadata as a product in their own right. As an output, the data and metadata from the surveys are distinct from the work carried out to collect data and carry out the surveys in the first place, themselves a large undertaking involving a range of members of staff and skills. Now that these resources are published and publicly available, they enable potential repeats of surveys, transparency, accessibility, data use and re-use.

1.5.2 Adding to knowledge in a way that has not been done before

Importantly, my work 'adds to knowledge in a way that has not been done before' (Gill and Dolan, 2015). This is on two levels – firstly by providing publicly accessible, quality assured data and metadata that can be analysed to look at change and past trends, which is effectively 'new' data. Rescued and fully documented data sets have enabled the prospect of novel analyses to be undertaken to answer particular questions, of the sort demonstrated in the paper regarding the characteristics of upland habitats of Britain (Paper 7), and to some extent demonstrated in the other six papers. Secondly, the approach to data rescue and publication was also novel and is set into the context of the movement towards open data and open access in science. When this work was begun, practical solutions and blueprints for achieving this were rudimentary. At the UK Centre for Ecology & Hydrology, the establishment and growth of the Natural Environment Research Council (NERC) Environment Information Data Centre has run in parallel with the body of work described, and provided the goal of ensuring the data in question could be stored safely for long-term use. The earlier papers and data deposits contributed to helping inform and direct best practice within the data centre to some extent, in terms of setting out requirements for metadata, supporting documentation and deposit procedures.

In summary, the work demonstrates how thorough and determined research and investigation can transform historically significant information, at severe risk of loss, into a publicly accessible, re-usable resource. The process of doing this has utilised, and contributed to the development of, modern technology and data repository procedures. It also demonstrates the lessons that can be learned in the future in terms of data capture and management, as well as approaches to other data rescue efforts and resources for future investigations of land use change.

2. Long-term monitoring: legacy data rescue and re-use

2.1 Ecological long-term monitoring

Ecological long-term monitoring can be defined in different ways. A useful definition is provided by Lindenmayer et al. (2012) as 'the systematic and regular collection of field data from a particular site or set of sites for more than 10 years'. In ecology, long-term studies are valuable as they can provide evidence for studying and understanding change in complicated systems. High-quality ecological information collected over long periods yields valuable insights into changes in ecosystem structure, key ecological processes and the services provided by ecosystems (Lindenmayer et al., 2012), answering questions relating to a range of issues such as climate change, ecological disturbances and biodiversity loss. Also, long-term investigations have provided critical data on a number of issues that are of concern to society (Elliott, 1990). Information on past trends is essential to inform future predictions and underpin attribution needed to drive policy responses, and long-term ecological data are often essential for quantifying ecological responses to environmental change such as natural, human or experimental disturbance (Carpenter et al., 1995; Likens, 1985). In terms of the 'Bunce' data, the value of repeat monitoring surveys has been proven already in the case of the Countryside Survey (Norton et al., 2012b) and the Woodland Survey (Kirby et al., 2005). A specific example of the value of the Countryside Survey data is the change in policy brought about by the quantification of hedgerow loss in Britain throughout the 1980s (Barr & Parr, 1994; The Hedgerows Regulations, 1997).

Long-term data sets can be used for multiple purposes, and in many cases have helped answer questions that their founders never considered (Magurran et al., 2010; Sheail & Bunce, 2003). Many were often initiated to answer specific questions, such as the Park Grass Experiment at Rothamsted in Southern England, now the longest running ecological experiment in the world, founded in 1856 to examine the effect of fertilisers on yield in hay meadows (Silvertown et al., 2006). The additional ecological value of the data collected at Park Grass soon became clear and they have been used to tackle problems ranging from the evolution of adaptation at a local scale, to the link between community composition and climatic perturbation (Kettlewell et al., 2006; Magurran et al., 2010; Silvertown et al., 1994). Data from Park Grass have been used to calibrate the Ellenberg values to assist in the interpretation of long-term change of Countryside Survey (Hill et al., 1999). Similarly, the Continuous Plankton Recorder began in 1925 with the goal of mapping oceanic plankton and relating them to fisheries, but has proved invaluable in addressing many questions, including community responses to ocean warming (Hawkins et al., 2013; Southward et al., 2005). The value of long-term data sets has been proven in a range of other examples, such as the North American breeding bird survey (Bystrak, 1981), the UK Environmental Change Network (Rennie et al., 2020), experimental plots at Wageningen (Pierik et al., 2011) and Cedar Creek (Symstad et al., 2003) and the Long Term Ecological Research (LTER) Network (Turner et al., 2003).

In the discipline of ecology, long-term monitoring is relatively rare and has often depended on chance rather than planning (for example Peterken and Backmeroff (1988)). Long-term monitoring is rare because it is difficult in comparison to short-term studies. It needs continued commitment of money, time, staff and facilities (Elliott, 1990) and hence is low priority for funding (Specht et al., 2018). The value of long-term data may not be immediately recognised due to a prolonged lag period until expression of key trends (Likens, 1992; Lindenmayer et al., 2012) and it needs constant analysis and scrutiny in order to keep it interesting (Elliott, 1990). As monitoring can be expensive, funding can prove difficult to secure, especially when the immediate value of data collection is not apparent. There are many areas where long-term monitoring can fail, including poor experimental design, bad data management, bureaucracy and poor leadership (Likens & Lindenmayer, 2018).

As stated by Lindenmayer et al. (2012), there are very few institutionalised and openly accessible ecological long-term studies; most are ad-hoc, fragmented and have a tenuous existence. Although growing awareness concerning ecosystem change has underlined the importance of long-term data sets, there are still relatively few biodiversity time series that span decades (Elliott, 1990; Magurran et al., 2010). The earliest ecologists showed great foresight when they initiated ecological experiments and began systematic data collections (Lindenmayer & Likens, 2010; Lindenmayer et al., 2012; Magurran et al., 2010), as in the case of the UKCEH 'Bunce' data sets (Figure 2.1). As stated by Professor Bunce, 'I always thought ahead but few other people did - hence the rarity of long-term monitoring' (R.G.H. Bunce, personal communication, 30/4/2020).

In the case of the early 'Bunce' surveys, rescuing the data ensures that the initial investment is not lost. They are rare examples of ecological data from a time of new methods and ground-breaking statistical methods. If further funding could be secured, added value could be gained from the surveys yet to be repeated (for example in Shetland, and 'Key Habitats' in England).

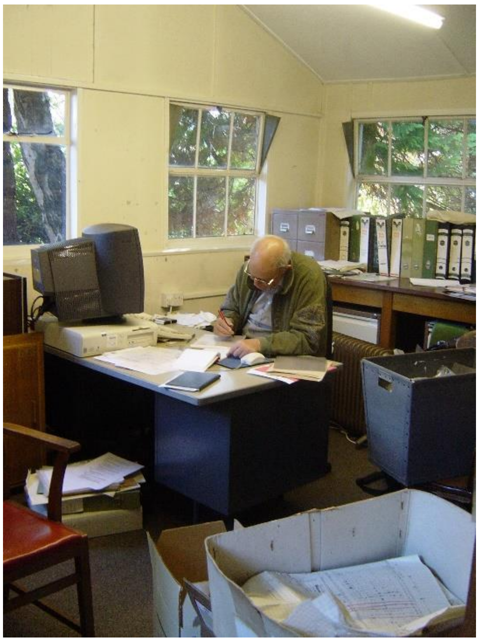


Figure 2.1. Professor Bunce at work at Merlewood Research Station

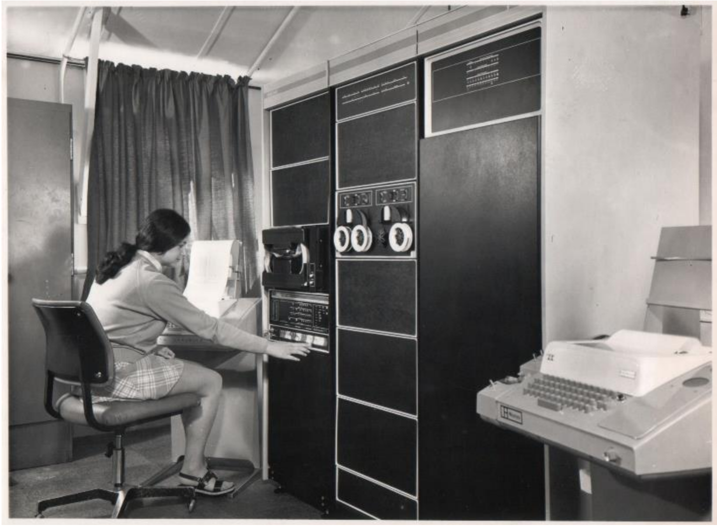


Figure 2.2. Contemporary data storage did not last the test of time.
Merlewood PDP8 Computer, c.1970.

Many major surveys cannot be repeated because there is not enough information regarding the original site locations or methods. Often methods were not standardised or specific, and depended on the subjectivity of the surveyors. For example, Hearn et al. (2011) have shown that most phytosociological surveys cannot be repeated, and Cherrill and McClean (1999) that the Phase I habitat surveys of the former Nature Conservancy Council are also not repeatable in a way that would allow a comparison of results across two separate surveys.

Whilst the first UKCEH 'Bunce' surveys were designed with the possibility of a repeat taking place, little practical thought was initially put into considering the repeat surveys. With hindsight, it is easy to look back and take the continuity and repeatability of Countryside Survey for granted as it has developed over time. However, in the 1970s, things were very different. Contemporary digital storage solutions, such as the paper tapes used by the 1970s PDP8 computer, did not stand the test of time (Figure 2.2), leaving the paper archives as the best and most vital resource for creating data products and documentation for the earlier surveys. As stated by Professor Bunce, 'Thought was given to storage but it did not last the test of time - just as well I kept a hard copy as the technology did not last' (R.G.H. Bunce, personal communication, 30/4/2020).

A lack of proper data management, including time for database development, data entry, data validation, analysis, interpretation and reporting of data can be one of the reasons why long-term monitoring programs may fail (Caughlan & Oakley, 2001) as well as the use of methods that are not standardised.

2.2 Data rescue

In 2013, Gibney and Van Noorden (2013) stated that data were being lost to science at a rapid rate, estimating that 80% of all ecological data used in papers dating back, even as recently as to the 1990s, had been lost. Although the situation has improved since 2013, this is still broadly the case. 'Treasure troves of data, and the knowledge they could offer, are left mouldering on shelves' (Griffin, 2017). A prime example of data from a large programme being lost is the International Biological Programme (IBP) (Worthington, 1965), which has failed to have a lasting legacy (Hampton et al., 2013; Michener et al., 1997). Set up in 1964 (and running until 1974), the programme planned to embrace large collaborative projects. However, this proved to be too much of a challenge and instead, many local, small scale projects were undertaken (for example at Merlewood (Bunce, 1968; Heal, 1968)). Whilst valuable science resulted from the programme, little or no raw data is currently available. Indeed, data and samples relating to the IBP work undertaken by the Nature Conservancy at Meathop Woods in Cumbria, and Moor House National Nature Reserve are stored in the archives at UKCEH, with no current plans for rescue.

The term 'data rescue' can be described as the ongoing process of preserving data at risk of being lost due to deterioration of the storage medium, and digitising current and past data into computer compatible form for easy access (Diwakar et al., 2008). All archives are unique and at risk of being damaged without any back-up of their holdings (Wilkinson et al., 2019). The value of rescued data is becoming increasingly clear in many disciplines, in order to answer questions regarding past and future changes and trends. The rescue of historic data can enable studies that would not otherwise be possible (Downs & Chen, 2017; Lindenmayer et al., 2012) and there are now many instances where historic data have proved useful. For example, logs recorded during past ship voyages have provided data for studying current weather patterns (Williamson, 2016), photos of glaciers from the past have yielded evidence of climate change (Rapp, 1996) and in the medical field, records on punch cards from the late 1950s, decoded decades later, have helped to show how varying levels of cholesterol predict later disease (Griffin, 2017; Pienta & Lyle, 2018).

One particular field where rescued data have been useful is the area of marine science. The Kenya-Belgium cooperation in marine sciences (KBP) project has been successful in recovering data from theses and reports resulting from marine and coastal research activities in the Eastern African region conducted between 1984 and 1999 providing a valuable resource to the scientific community (Knockaert et al., 2019).

Data have been collected in the English Channel since the 1880s by the Marine Biological Association and associated bodies (Southward et al., 2005), and a good example of using old data to make future predictions has been the use of the Southward barnacle time-series (1950s–1987). Climate-driven fluctuations in abundance of southern warm-water *Chthamalus spp*., and northern cold-water *Semibalanus balanoides* have been modelled to understand the role of competition in modulating climate responses. This has enabled better predictive modelling of future situations under different climate change scenarios (Hawkins et al., 2013). Caldwell (2012) also describes how tide gauge data have been rescued in order to investigate global sea level change.

Progress has also been made with soil data, for example under the GlobalSoilMap project (Arrouays et al., 2017) which aims to rescue and collate soil data sets from around the world. Weather and climate data are also areas that have featured in data rescue initiatives (Diwakar et al., 2008; Williamson, 2016).

Data rescue initiatives may be straightforward, for example, simple automated digitisation and/or data entry from weather charts or ship's logs by citizen scientists (for example, Indian Weather Charts (Diwakar et al., 2008) and the Rainfall Rescue Project (BBC, 2020)). However, in the case of the UKCEH 'Bunce' monitoring data sets, it is important to emphasise the complexity of the data, which represent large-scale ecological surveys with a range of different components, spatial data and complex data collection protocols. In the case of complex data such as these, data rescue needs to occur before the data in question become completely inaccessible or unusable, and ideally should occur while those scientists or others familiar with the data are still available to provide important information about the data, their origin, collection, and management, and quality (Downs & Chen, 2017; Lindenmayer et al., 2012).

As old storage media become unusable, accidents such as fire, flood and mistaken discard occur, and the original data collectors are no longer available to question, the opportunities for rescuing many data sets are fading (Griffin, 2017; Whitlock, 2011). In light of these concerns, in the last 10-15 years, NERC has placed an increasing emphasis on good data management practices in relation to grant-holders and its institutes (including UKCEH). Whilst the current emphasis within UKCEH is upon planning data management tasks in order to avoid the loss of newly collected data in the future, the availability of a small amount of resource and new technical capabilities for cataloguing and publishing data (via the Environmental Information Data Centre) have meant that the UKCEH 'Bunce' data sets have been able to be secured and have now been preserved before risking loss.

2.3 Open access to scientific data and publication

2.3.1 Concept of open data

Once a set of data has been rescued, there are many advantages to making it generally available or 'open', in order to gain maximum value from the data, and hopefully, ensure longevity. Open data can be defined as 'data that's available to everyone to access, use and share' (Open Data Institute, 2017). To be strictly 'open', data should be completely free, with no restrictions (Attard et al., 2015). However, this may not always be practical or possible.

There are many benefits in sharing data. Advantages include the ability to find and re-use data which have already been collected, in order to save time and money by avoiding redundant data collection. Data sharing also enhances opportunities for collaboration and meta-analysis, and for answering questions not previously posed by the original data collectors. As ecology is an integrative, collaborative discipline, there is an obvious need for open access to data (Arrouays et al., 2017). Much current work tends to integrate a range of data sets to answer a wide range of questions (for example, the Countryside Survey Integrated Assessment (Smart et al., 2010a) and Ecosystem Services and Natural Capital (Henrys et al., 2015; Norton et al., 2012a)) with now more of a focus on large scale, long-term data sets than in the early days of the discipline.

The concept of open data has been around since the 1950s (arising in preparation for the International Geophysical Year (IGY) in 1957). However, until the rise of the internet, this was difficult to achieve in practice. Open data movements (in relation to publicly funded data, and a drive for enhanced transparency in government) began to proliferate in the last 20 years with, for example, the Public Sector Information (PSI) Directive in Europe, the Obama Open Data Initiative in 2009 in the US, the Open Government Partnership in 2011 (multinational) and the G8 Open Charter in 2013. Open government data portals resulting from such movements (such as data.gov.uk) provide a means for dissemination of the open data (Attard et al., 2015).

Data repositories and centres, such as the NERC Environmental Information Data Centre (EIDC) (established in ~2012) have been multiplying rapidly in the last 10-20 years (see Michener (2015) for examples). In terms of scientific data, NERC alone hosts a network of data centres including those covering atmospheric science and earth observation (CEDA, 2020), oceanographic science (British Oceanographic Data Centre, 2020), polar data (UK Polar Data Centre, 2020) and geoscientific data (National Geoscience Data Centre, 2020). These are not all as recent as the EIDC. For example, the British Oceanographic Data Centre has its origins 1969 as the British Oceanographic Data Service (BODS) and was a pioneer in its approach to (marine) data management. EIDC itself was largely born out of a requirement to improve the historically poor approaches to data management practices generally within UKCEH, and a recognition that long-term projects, like Countryside Survey, needed a long-term platform for data curation to withstand the turnover of staff (J.W. Watkins (Head of EIDC), personal communication, 29/9/2020).

A data centre at a basic level includes a technical infrastructure for data storage, plus ongoing support, management and guidance for best practice (Collins, 2011). The concept of data centres can again be dated back to the IGY in 1957. In planning for the IGY, recommendations by the planning committee were made in 1955 mentioning that data centres should be prepared to handle data in machine-readable form, which at that time meant punched cards and punched tape (Ruttenberg & Rishbeth, 1994). 'Although existing data centres have varying standards of quality and tend to be discipline specific, they hold tremendous promise for increasing the scope, coverage and societal relevance of ecological and biodiversity studies' (Berkley et al., 2009). Nevertheless, the data in these repositories, even ten years ago, were not representing a significant proportion of the extensive ecological, environment and biodiversity data collected each year (Berkley et al., 2009; Whitlock, 2011), not even considering legacy data sets. The situation has improved thanks to changes in data policies from funders and journals. In the previous 5-10 years, many scientific journals now have a policy in place requiring authors to deposit their supporting data with a data repository. Publishing the UKCEH 'Bunce' data sets has been beneficial to colleagues at UKCEH, in enabling them to meet these requirements. As stated by the Head of EIDC, 'Being able to demonstrate case studies of published data sets, such as the Bunce Woodland Survey, makes big in-roads in impressing potential funders of projects' (J.W. Watkins, personal communication, 29/9/2020).

In UKCEH, being largely publicly funded, a sizeable proportion of data generated must conform to the data policy of NERC. In 2011, this policy was updated in support of the government's developing agenda on open access to public data, including the stipulation that 'for all research publications produced by NERC's own staff, the supporting data will be made available through the NERC data centres'. Central to the policy is that NERC-funded scientists must make their data (of long-term value) openly available within two years of collection and deposit it in a NERC data centre for long-term preservation. The aim is that all NERC-funded data are managed and made available for the long-term for anybody to use without any restrictions (NERC, 2020). In terms of publicly funded data, it is now widely expected that scientific research should be demonstrated to be transparent and reproducible, to prove its value to the taxpayer (UKRI, 2020).

In addition to an emphasis on improving digital data procedures and policies, the focus within NERC, and particularly within UKCEH, on data and information being recognised as an asset has also extended to libraries, and more specifically archives. It is fortunate that UKCEH has been able to store historic information since the 'Bunce' surveys occurred, for example on old network drives, databases and in document boxes. However, this has really been through the foresight of particular individuals rather than planned management. UKCEH has recognised that their approach to archiving data and documents has not been adequate in the past and therefore in the past decade, UKCEH has improved procedures for physical archives, safeguarding information for the future. This type of archived information, such as maps, photos, reports and handbooks, is essential for data rescue to be carried out successfully, in order to provide adequate information to describe a related data asset. The controlled storage of 'paper' data sets also raises the prospect of other successful data rescue projects.

2.4 Requirements for the re-use of rescued data

In order to be able to share, and consequently re-use, data sets effectively, there are a range of basic minimum requirements concerning the quality of the data and metadata in question. Ecological data are unusually complex (Zimmerman, 2008) and present a particular challenge for management and preservation because they are geographically, taxonomically and temporally unique (Ellison, 2010), and also heterogeneous (Reichman et al., 2011; Wieczorek et al., 2012). This is illustrated by the example of the Countryside Survey in Figure 2.3., and certainly applies to all of the rescued UKCEH 'Bunce' data sets, which incorporate a range of time-spans, landscape types, nomenclatures, themes and methods. Rendering the data sets usable and publishable has required much technical work and research involving the consultation of primary sources.

It is important to emphasise that the availability of research data is not the same as existence of fit-for-re-use data (Tani et al., 2013). In order for potential users of data to re-use third party data confidently, they must have some indication of the quality of both the data and the relevant metadata; the standards are higher for data that needs to be repeated in long-term studies (Specht et al., 2018). Specific metadata must be provided to researchers so they can understand the data being supplied, and evaluate their suitability.

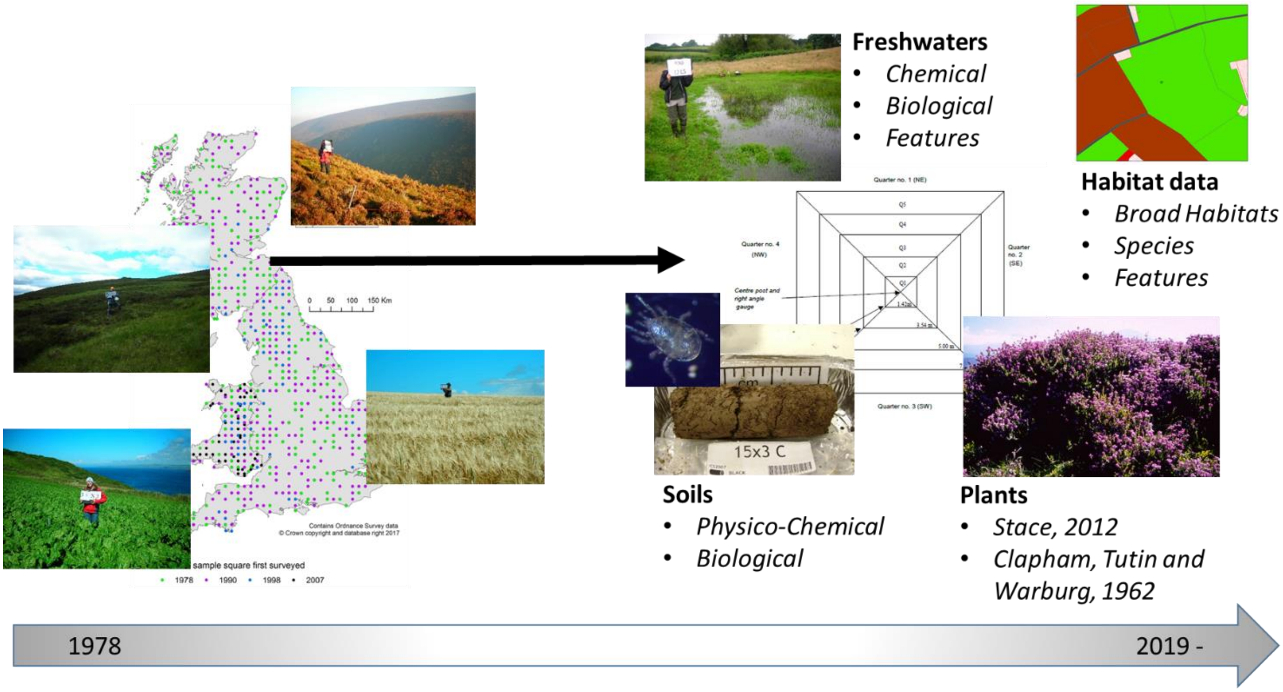


Figure 2.3. Ecological data are geographically, taxonomically and temporally unique – as illustrated by examples from the Countryside Survey

2.4.1 Data quality

In terms of data sets, there are many areas where quality can be assured, and a wide range of definitions of 'data quality' (Guptill & Morrison, 2013; Michener & Brunt, 2009). The key outcome for a published data set is that it is fit for use/re-use.

For all of the UKCEH 'Bunce' surveys, a range of controls were involved in the data collection. During all of the surveys, comprehensive training courses were held to ensure constancy of recording, and all survey teams were initially accompanied by a supervisor. Regular visits into the field were made by project leaders to ensure consistency and quality in data recording according to criteria laid out in the field handbooks. Repeatability in terms of relocating plots has been achieved by the use of detailed plot maps, measurements and photographs, sometimes in addition to buried metal plates enabling surveyors to re-find plots with a metal detector. Since 1998, these methods have been complemented by global positioning systems.

Despite careful quality control when collecting data, many additional quality checks on the digitised data had to be carried out. A summary of these is provided here, however a detailed description of these may be found in Section 3 and Appendix ii.

EIDC has controls on the quality of data that can be accepted for ingestion (EIDC, 2020b). This guidance concerns file formats, naming and coding conventions and data structure. As early deposits into the data centre, lessons learned from the UKCEH 'Bunce' monitoring data sets partly helped to inform some of this guidance. Lawrence et al. (2011) outline a useful framework of quality checks, into which those undertaken can be categorised:

• ***Is the format acceptable?*** Data (and also metadata) stored for the long-term has to be as 'future-proofed' as possible. EIDC recommend comma separated value (.csv) files as the preferred format for depositing tabular data sets as they are proven to be robust and future-proof, allowing reading and viewing of the data through a wide variety of common software tools and conversion to many common formats (EIDC, 2020c). All of the UKCEH 'Bunce' tabular data sets are published as .csv files, with spatial files being available as ESRI shapefiles.

• ***Are data values internally consistent?*** For example, standard validation checks included plot and site counts to ensure no duplicate numbering and hence double counting of plots. Furthermore, range checks were undertaken where possible, for values falling within expected ranges, such as soil pH or slope values. Values were updated to ensure consistency (for example update 'Birch', 'Betula', 'BIRCH' to '*Betula sp.*')

• ***Does the data represent reality with sufficient accuracy to use? Is the data of tolerable precision?*** In the case of plot locations, plots were marked on 1:10,000 maps in the field and later digitised. These locations have been proven to adequately re-locate plots in subsequent surveys (Wood et al., 2015).

• ***Does the extent and coverage of the data match expectations?*** Checks that all sites and plots recorded are correctly included in the data sets, as well as the items mentioned as being recorded in the handbook are included in the data.

• ***Are the data values reported physically possible and plausible?*** Checks were undertaken against code lists, and scanned for values that would be in the expected range.

• ***Is the data validated against an independent data set?*** As the 'Bunce' data sets are unique, this is not easy to achieve. Taking the example of soil pH values from the Woodland Survey, they were analysed using quality control measures as outlined in Allen (1989). These included the analyses of certified standard reference samples within batches. The 'Key Habitat' strata were also validated against other data sets at the time of their initial creation (Wood et al., 2018).

2.4.2 Metadata quality

Metadata may be defined as 'information about data' - i.e. the information required to understand data, including data set contents, context, quality, structure, and accessibility (Michener et al., 1997). Metadata describe the 'who, what, when, where, and how' about every aspect of the data (Michener, 2006).

As integrated studies are becoming common in ecology, ecologists are increasingly using data that have been collected by other scientists from numerous disciplines other than ecology, to answer questions such as those relating to biodiversity loss or climate change (Reichman et al., 2011; Zimmerman, 2008). Most scientists recognise that incomplete and inadequate metadata have been a significant technical barrier to data integration and analysis efforts (Michener, 2006). When considering data re-use, few, if any, ecological data sets are perfect or intuitive (Michener et al., 1997) and metadata published with papers is often not good enough (Kervin et al., 2013). Hence, it is likely that highly detailed instructions or documentation will be required for scientists to accurately interpret and analyse historic or long-term data sets, as well as data resulting from unfamiliar research or complicated experimental designs (Michener et al., 1997). As Professor Bunce states, 'Preserving data sets with notes is vital – they must be understandable by the next generation'. (R.G.H. Bunce, personal communication, 30/4/2020).

The most important reason to invest time and energy in developing metadata is that human memory is short. Many processes can lead to the loss of information through time (Fig. 2.4). Some of these processes operate continuously, such as the gradual degradation of storage media containing the data, whereas others can be categorised as discrete events, such as the retirement or death of the scientist who collected the data, obsolescence of storage technology, or the loss of storage media through catastrophic events. Although loss of metadata can occur throughout the period of data collection, the rate of loss is likely to increase after project results have been published or the study has been terminated (Downs & Chen, 2017; Michener et al., 1997).

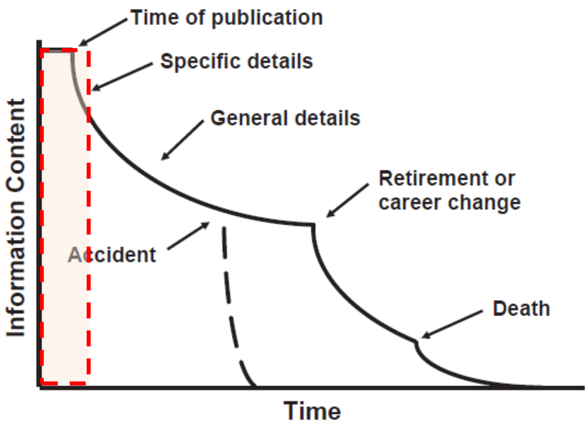


Figure 2.4. Illustration of the natural degradation in information content associated with data and metadata – information entropy (Michener, 2006). Area highlighted in red is the suggested ideal window in which metadata should be captured in order for data to still be useful for re-use (i.e. as soon as practically possible).

As funding for long-term monitoring is usually uncertain, it is essential for experiments to be properly documented, in the hope they can be repeated at a point in the future where funding may again be available; 'data longevity is roughly proportional to metadata comprehensiveness' (Bowser, 1986). The rescue of the early UKCEH 'Bunce' data sets has coincided with increased attention in the last decade to the development of high-quality data sets and their metadata within UKCEH, and thus provide a blueprint for achieving this, to a certain extent contributing to the approach UKCEH has taken in developing policies and procedures for the NERC Environmental Information Data Centre. Both data and metadata must meet certain standards of quality, and ecologists have been working to establish metadata protocols to standardise written information about data to sets to make possible data sharing and re-use (Jones et al., 2001; Michener et al., 1997). Identifying the essential information for a particular data set is difficult, and it has been acknowledged that 'there is no unique, minimal and sufficient set of metadata for any given data set since sufficiency depends on the use to which the data are put' (Zimmerman, 2008).

Metadata for a given data set can be broadly split into two groups, as distinguished by Bretherton and Singley (1994) as 'guide' and 'structural/control' metadata, although other sub-divisions have been suggested, for example Lawrence et al. (2011); Michener et al. (1997). 'Guide' (or discovery) metadata is the basic information used to describe a data set, enabling users to find the data set and determine its usability. 'Structural/control' (or field level) metadata consists of a detailed collection of information required for data re-use. A more detailed description of these may be found in Section 3.

Whilst there is a range of standards regarding discovery metadata (see section 3), for field level metadata, establishing an acceptable level of quality can be more subjective on account of the large range of information that may, or may not, be included. This is particularly true of ecological data which cover a breadth of different collection methods, data types, spatial and temporal ranges. The EIDC has gradually developed a set of guidance regarding (field level) metadata quality, partly informed using the experience from the UKCEH 'Bunce' surveys. As stated by the Head of EIDC, 'An area of strength of the EIDC has been the development of guidance regarding supporting information and metadata accompanying ecological data deposits' (J.W. Watkins, personal communication, 29/9/2020).

The current guidance provided by EIDC (presented in Appendix v) suggests a number of headings by which to arrange the metadata, such as 'Experimental design/sampling regime', 'Nature and units of recorded values', 'Quality control' and 'Details of data structure'. In order to ascertain these details in relation to the UKCEH 'Bunce' data sets, much research was undertaken among primary sources (see Section 3). Again, Lawrence et al. (2011) suggest a framework addressing metadata quality issues. Those relevant to the UKCEH 'Bunce' data sets are described below:

- ***Is there sufficient quality metadata describing the format and physical content?*** Metadata describe the format of the data files, mostly comma separated value (.csv) files in the case of the UKCEH 'Bunce' data sets, and also the content, including descriptions of each column name (for example 'SITE' = Survey site name)
- ***Is there sufficient quality metadata describing provenance and context? Has the data changed in some way since it was measured? Is the processing chain visible and well documented? Have all the human interactions with the data prior to ingest/publication been recorded?*** For example, what methods were used to collect the data or analyse soil samples? Who collected the information, and is there evidence of their reliability?
- ***Is there existing metadata (or are there references) already making assertions about the quality and usefulness of the data? If so, are these included in the metadata?*** References to previously published reports or papers may have information regarding the quality of the data.
- ***Is there suitable quality discovery metadata*?** ***Are all the available metadata conforming to standards?*** As standards exist for discovery metadata, this can easily

be ascertained by depositing data in a reputable data centre such as EIDC, which uses recognisable standards (based on UK GEMINI 2.3 (AGI, 2020a)) to catalogue records.

- ***Does the metadata use appropriate, controlled vocabularies?*** Are there specific ways of organising knowledge? For example UKCEH have controlled vocabularies for chemical determinands (UKCEH, 2020a).
- ***Is there an existing user community? Is that community happy that the data are usable?*** This can be achieved by publishing the data in a peer reviewed publication, as per the UKCEH 'Bunce' data sets.

More specific examples are given in Section 3, and in Appendix iv.

2.5 Challenges overcome in publishing the UKCEH 'Bunce' data sets

There is a range of reasons as to why the earlier 'Bunce' data sets have never been publicly available in a re-usable form until now, and why the methods and data have never been documented before. Many of the reasons are widely applicable to other comparable legacy data sets, and include issues such as accessibility, 'findability', compatibility and understandability, as comprehensively documented by Zuiderwijk et al. (2012). These issues have developed into the 'FAIR Guiding Principles for scientific data management and stewardship' (Wilkinson et al., 2016): data should be Findable, Accessible, Interoperable and Re-usable.

2.5.1 Technology and technical challenges

A widespread issue concerns problems relating to physical data storage media, due to the pace at which they are liable to become obsolete. Since computers became commercially available, data have been stored on media ranging from punched cards and paper tapes, to a range of sizes of magnetic and compact discs (Figure 2.5). Unless the data are constantly upgraded to the latest formats and storage methods, they will quickly become unusable and may be completely lost (unless they have a secondary source, perhaps on paper). However, it is also commonly accepted that data curation is a difficult job, and most data producing scientists have neither the time nor the inclination to focus on it (Callaghan et al., 2012). Much of the earlier series of 'Bunce' data were stored on tapes and disks that are now obsolete, and would have been lost if hard copies had not been kept. Depositing data in a reputable data centre for the long-term should take care of this issue, as it becomes the responsibility of the data centre to ensure the ongoing usability of the data they curate.

Figure 2.5. Data storage: Merlewood Computer in the 1980s (top left), paper computer tape c.1970 (bottom left), field survey documents in archive boxes at UKCEH (right)

In terms of data management, at the time the 'Bunce' surveys were initiated, technologies were limited in terms of the capabilities for producing data products in a robust, digital format that could be stored for the long-term. The earliest UKCEH 'Bunce' surveys were carried out using paper recording sheets and initially data were stored and analysed using paper computer tapes (see Figure 2.5). Whilst later surveys (for example, the 'Key Habitats' survey) did make use of more technologically advanced tools evolving at the time (for example Geographical Information Systems and corporate databases such as Oracle), software and infrastructures for managing spatial data are constantly improving in terms of functionality and ease of use, and knowledge regarding the most durable formats for storing data for the long-term have advanced considerably. For long-term electronic storage, some file formats are preferable to others for preserving longevity. The Digital Curation Centre (DCC) was launched in 2004 with support from Jisc (Jisc, 2020) and the Engineering and Physical Sciences Research Council (EPSRC) and provides guidance for best practice in relation to the optimal file formats for long-term storage and re-use (DCC, 2020). Preferable formats are open source rather than proprietary, in common use and least likely to degrade. The Environmental Information Data Centre also has guidance concerning the data sets they are willing to store (EIDC, 2020c), largely following these recommendations.

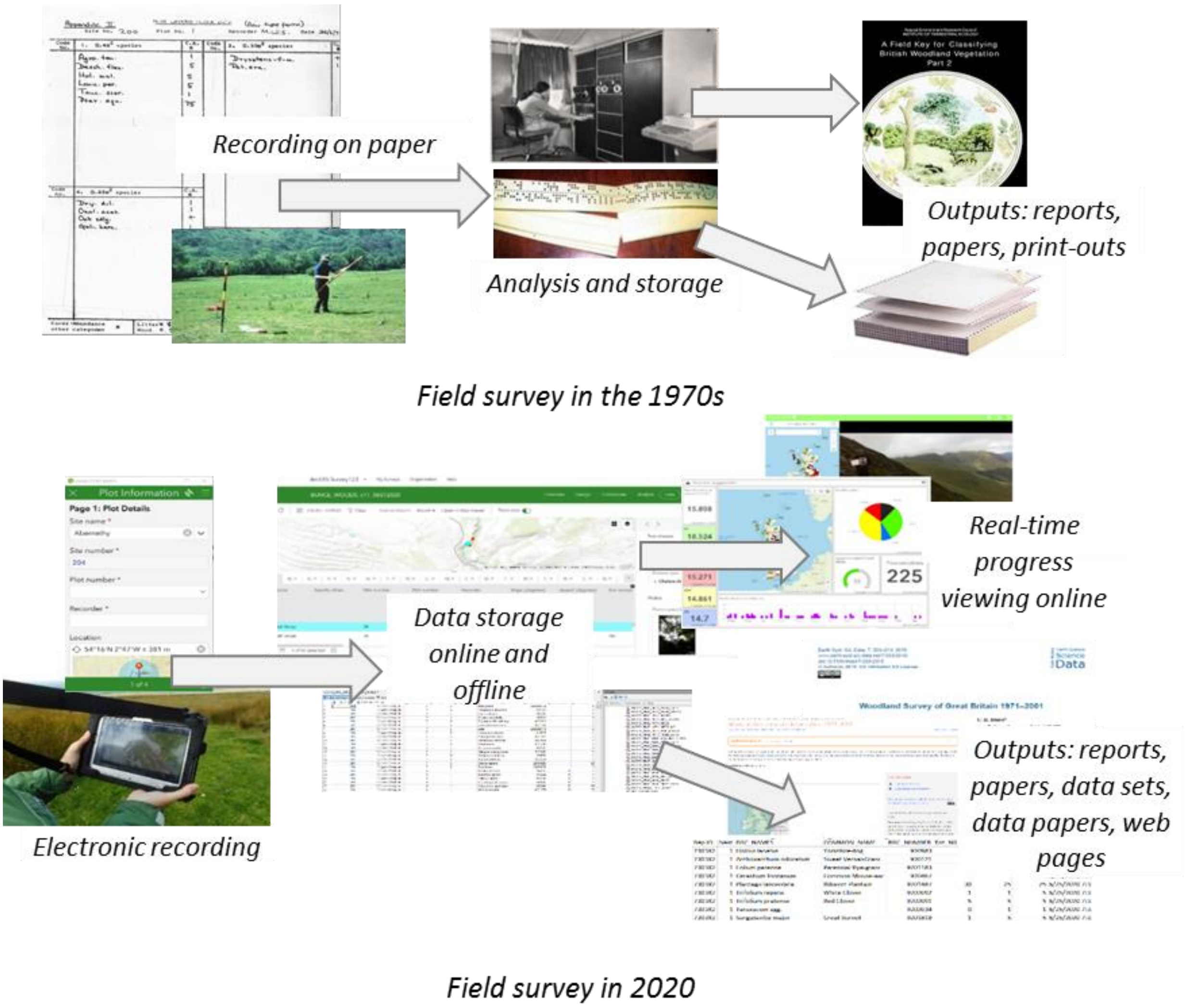


Figure 2.6. Comparison of field survey techniques, 1970 and 2020

In the past, data entry from paper recording sheets was time-consuming and was a potential source of the introduction of error. Since the 2007 Countryside Survey, electronic data capture systems have become the standard method for most of the large surveys undertaken at UKCEH. Surveys undertaken since 2019 have started to make use of ESRI's ArcGIS Online tools, as described in a recent blog (Wood, 2019). New ArcGIS tools (in particular Survey123 and Collector) have brought a range of benefits to data collection. In terms of software, the apps are small and easy to install, compared to the time-consuming installation of Windows desktop software used previously. They can also be used on small and lightweight devices. For surveyors, data are easily submitted straight into the central database online, including site locations, photos and edited maps, making the need for an additional camera and associated photo transfer redundant, and minimising post-processing for office staff. Soil samples can be tracked using barcoding in the field, helping laboratory staff manage and process samples more efficiently. In terms of monitoring the survey, staff back in the office can supervise how the teams are progressing in the field via the Operations Dashboard and Photo Viewer apps (Figure 2.6).

2.5.2 Skills and Resources

In some cases, data can readily be rescued, or at least digitised, fairly easily in an automated way (for example automated image processing of historical weather charts (Diwakar et al., 2008)). However, the task of unravelling a set of coded ecological data and methods from up to 40 or 50 years ago is more complex, and requires specific technical skills in determining the best methods of organising data, then implementing and documenting this data design. Other essential skills include the ability to develop protocols for data processing and extracting data from databases, the ability to clearly document the data and associated information, and train colleagues and other users in how to extract the data they require. An interest in the subject matter is also a useful asset, along with the enthusiasm to perform the detective work needed to understand the historic data sets and present it in such a way that will make sense to potential future users of the data.

In past decades, the importance of curating historic data was not recognised. Also, neither staff, time nor appropriate resources were available to undertake such a large task in one funded project. The teams of staff who collected the data had different skills to those required in order to manage and curate data (with an emphasis on botanical knowledge, rather than computing and data management skills). The specific technical challenges regarding the UKCEH 'Bunce' data sets are described in detail in Section 3 and Appendix ii.

2.5.3 Lack of a wider framework for long-term data storage

In practical terms, the idea of public data archives and repositories to deposit (ecological) data and metadata is a relatively new one. Before these repositories were in place, there was nowhere to store rescued data long-term (Michener, 2006). Thanks to advances in technology, practical solutions regarding long-term data storage and security have been developed in the last decade (such as the inauguration of EIDC in ~2012). There is also a wider appreciation of the importance of a need for careful data management planning when undertaking scientific work. This is particularly the case within nationally funded institutes such as UKCEH, and in relation to public funded grants, for example those provided by NERC. At UKCEH, in the last decade, there has been an increased emphasis on improved data management procedures and protocols, and the provision of staff forming an Informatics Liaison Network to assist scientists with these issues. In many cases, a careful data management plan must be submitted at the time a project is proposed (for example, UKRI-NERC (2020)). A data management plan helps scientists consider issues such as the type and amount of data to be collected, the long-term plan for storage, metadata and documentation, licensing and copyright before a project starts.

As discussed above, the idea of open data is quite new in terms of practical applications, and certainly since the birth of the internet. Until the creation of data centres such as the NERC EIDC to provide a long-term storage solution, there was no goal or incentive to work on data that would potentially become at risk of loss (Costello et al., 2013; Molloy, 2011). The ability to deposit the data in long-term storage provides a sense of completion to the task of rescue.

2.5.4 Privacy and data protection

Site locations of the 'Bunce' UKCEH surveys are held confidentially, most notably in the case of the Countryside Surveys. The surveys are largely carried out on private land and due to the statistical design of the surveys, it is important to avoid the introduction of bias into the sample. Therefore, preserving the representativeness of sampling sites and the goodwill of landowners are both essential elements to the future of surveys to ensure the scientific integrity of the sampling strategy, the protection of the environment, and help to ensure future permission from landowners to survey their land. It could be said that restricting access to the specific site locations raises a barrier to truly 'open data'. However, the reasons for the restrictions are valid, in order to avoid compromising future surveys, and the restrictions do not preclude the re-use of the data.

2.5.5 General attitudes

In general, attitudes surrounding data and data sharing are slowly changing. Given that the currency of academic credit has traditionally been based around journal publications, and the historic difficulties associated with publishing data, data sharing has generally been viewed with a variety of opinions from enthusiasm to scepticism or outright hostility (Costello et al., 2013; Molloy, 2011; Reichman et al., 2011). From personal experience, many scientists are protective over what they regard as 'their' data sets and there is often a general reluctance to share, especially where work is undertaken in a competitive environment. This attitude is slowly changing, with credit being duly given to scientists in the form of citations issued for data sets, peer-reviewed data papers and a general appreciation of the mutual benefits of sharing and collaboration. Perhaps the biggest change in attitudes has been brought about by new requirements brought in by research funders and publishers, and gradually scientists are realising they will not win grants or be able to publish their research without complying with new rules surrounding the publication and long-term storage of data (Collins, 2011).

2.6 Example uses/potential uses of UKCEH 'Bunce' monitoring data

Having all of the 'Bunce' data sets now published, well-documented and stored in a data repository has enabled new research to comply with new open access data policies required by many journals. The rescue work provides potential for a range of new analyses, and also the prospect of repeat surveys, thus also change analyses.

2.6.1 Woodlands re-surveys

The work completed on the data and metadata on the Woodlands and Scottish Pinewoods (Papers 1 and 2) has gone a long way to facilitating a re-survey, in collaboration with the Woodland Trust and other funding bodies, starting in 2020. Having the data in a re-usable, well-described format demonstrates the possibilities of analysing change, provides a structure for newly collected data, and clearly demonstrates to funders and data users what the potential of the rescued data might be. The new availability of re-usable, documented data now opens up the possibility of further work to a wider range of researchers. With the addition of new data from 2020-22, previous analyses can be repeated to explore further changes.

The Scottish Pinewoods survey has never been fully repeated, until the survey currently underway, therefore the availability of new data raises many new avenues of research. The broadleaved survey was repeated in the early 2000s and the data have proven their worth, having been analysed and used in a range of ways to answer a variety of questions. After the first survey, publications arising from the data included the production of 'A Field Key for Classifying British Woodland Vegetation' (Bunce, 1982, 1989). The survey was also described in Bunce and Shaw (1972) and used to put British woodlands into a European context in Bunce (1981). Following the second survey in 2001, a range of analyses were undertaken, as described in Kirby et al. (2005), focusing on changes that had taken place between the two surveys. Some of the conclusions from the main findings were that there had been an overall increase in soil pH, particularly in organic soils, but there was no increase in the mean level of soil organic matter. Most tree and shrub species remained stable in terms of their frequency of occurrence at plot and site levels, although 15 species (nine of these shrubs) declined, whilst five other species (four of them conifers) increased. There was a net loss of stems from the smallest size classes (particularly less than 10 cm DBH) with some smaller gains in the 30-60 cm classes. Stems greater than 60 cm remained scarce, although different species revealed distinct patterns of variation. Overall ground flora species richness declined by up to 32 % at a plot level (Kirby et al., 2005). Other work has included looking at the impact on species richness in woodlands after the extreme storm occurring in October 1987 (Smart et al., 2014), aggressive dominant species (Marrs et al., 2013) and the effects of landscape-scale environmental drivers on the vegetation composition of British woodlands (Corney et al., 2004).

2.6.2 Shetland Survey

The original Shetland survey was designed for monitoring change, and having the newly rescued data provides great potential for a repeat survey. The results from such a survey could help justify conservation effort, help target agri-environment schemes more effectively to support crofting but also bring conservation benefits and funding to support peatland restoration.

In Shetland, rapid social and economic changes have taken place since the UKCEH survey, driven by the arrival of Sullom Voe oil terminal (Hill et al., 1998), changes which also affected agriculture. There have been further changes over the last decades, including local urban development pressures, plans for extensive wind farm development (Viking Energy, 2016) and a decrease in peat cutting (the dominant energy source for many households in the 1970s), although this did start to increase again slightly in response to increased energy prices. In the future, climate change may affect Shetland's boreal and alpine vegetation, threatening species such as *Cerastium nigrescens, Arenaria norvegica spp norvegica, Saussurea alpina, Alchemilla alpina* and *Luzula spicata*.

2.6.3 Countryside Survey

The data from Countryside Survey were not at risk in the same way as the other data sets described, with the exception of certain data sets from the 1978 survey, on which the rescue work focused. However, having all of the data from Countryside Survey now published, well-documented and stored in a data repository has enabled new research to comply with new open access data policies required by many journals. In terms of the 'Key Habitat' work (Paper 6), the rescued data raise the prospect that the additional sites could be incorporated into future Countryside Surveys. The additional sites could provide data on rarer landscape and habitat types in order to add value to the standard Countryside Survey. Whilst funding may be hard to obtain in the current economic climate, the data are secure should a repeat possibly take place at some point in the future.

The paper regarding the ecology of British upland landscapes (Paper 7) demonstrates an example of how the UKCEH data can be used, thus demonstrating the value of this type of data and documentation. The data used in the analyses of upland landscapes, vegetation classes and vegetation species, were gathered during the Countryside Survey, thus providing a comprehensive, definitive set of statistics for the British uplands. The wealth of data available from this survey allows an analysis of British upland landscapes, arguably not possible from any other data source, and not originally one of the primary drivers for undertaking the survey and data collection in the first place. By extension, further work could be undertaken by analysing the data from the earlier surveys to compare the composition of the uplands through time and identify change.

There are other examples where the long-term data sets provided by Countryside Survey have been able to answer a range of questions, demonstrating how invaluable data of this nature can be. After 2007, the data contributed to many areas of the UK National Ecosystem Assessment (NEA) (Watson et al., 2011), which articulated ecological status and change in terms of ecosystem services (ES). Countryside Survey plot data have been used to produce maps of natural capital for policy makers (EIDC, 2020a) and to help in understanding the factors influencing spatial differences in ES delivery (Henrys et al., 2015; Norton et al., 2016). Countryside Survey data sets have also made a unique contribution to the development of plant species niche models for ecosystem dominants and many rare species in Britain (Henrys et al., 2013; Hill et al., 2017; Smart et al., 2010b). The statistically robust, national scale of the Countryside Survey vegetation data set makes it ideally suited to detect realistically scaled relationships between global change drivers, such as pollutant deposition (Maskell et al., 2010; Smart et al., 2005a; Stevens et al., 2016; van den Berg et al., 2016) as well as other drivers of eutrophication and land management change (Smart et al., 2005b; Smart et al., 2002; Smart et al., 2003; Smart et al., 2012; Smart et al., 2006). Habitat specific studies, such as those relating to woodlands (Kimberley et al., 2013; Kimberley et al., 2016; Petit et al., 2004) and hedgerows, (Critchley et al., 2013; Garbutt & Sparks, 2002; McCollin et al., 2000) have been facilitated through the use of Countryside Survey data.

Data have also been used to assess relationships between biodiversity in birds and plants, and habitat and landscape feature presence and extent (Rhodes et al., 2015; Smart et al., 2010b). Drivers of environmental change may be investigated, for example the effects of agricultural intensification (Petit et al., 2004) and farming practices (Potter & Lobley, 1996) on habitat quality and extent. The loss of hedgerows has been a key concern since the end of the Second World War, and Countryside Survey data have proved useful in determining the extent and nature of changes since 1984 (Barr & Gillespie, 2000; Barr et al., 1991; Norton et al., 2012b;

Petit et al., 2003) and applying these to policy changes (Barr & Parr, 1994). Countryside Survey data have contributed to determining policy, for example the Hedgerow Regulations (The Hedgerows Regulations, 1997).

Section 3. Technical Implementation

3.1 The generic approach to rescuing ecological survey data

Rescuing, reconstructing and documenting the data sets was a challenging process that can be described and presented in the following five steps:

Step 1: Identify available resources. This step involves identifying and locating the available resources relating to the data set in question. This includes paper documents such as field handbooks, field recording sheets, photographs, maps and code descriptions. In some cases, there may also be extant digitally recorded resources.

Steps 2 & 3: Process data sets and assemble metadata. These steps concern the construction of the data (Step 2) and metadata (Step 3) resources. At the most basic starting point, data may either have to be entered digitally, or it may be the case that some digital resources may already exist. At this stage, a sensible data schema must be designed (a plan for how the data and data elements will be named, structured and related – see below). The data must then be structured to fit into the planned schema. Whilst working with the data, questions are likely to arise, concerning the nature and units of recorded values, field methods and experimental design. Answering these questions with the assistance of available documents and the personal knowledge of the original scientist is essential in order to construct adequate metadata to accompany the data set. These two stages may be iterative, as organising the data relies on some knowledge of the metadata, but also constructing metadata also relies on knowledge of the data. At this stage, the data may be stored in an internal file storage system or database, potentially in a proprietary format (for example, the UKCEH data are stored in secure Oracle databases, accessible by scientists in a specific UKCEH research group, with metadata in a collaborative online storage system).

Step 4: Produce outputs. This concerns the creation and output of re-usable, publishable data products. This includes both the data set itself, presented in a robust, non-proprietary format, and the explanatory metadata in a comprehensive, user-friendly document.

Step 5: Publication. This stage relates to the publication of the data products created in step four, via the route of deposition in a data repository (in this case, EIDC), and the issuing of a Digital Object Identifier (DOI) and citation for the data set. The publication stage may also reference, analyse or describe the published data set in a peer-reviewed scientific or data journal.

The steps followed are outlined in Table 3.1 and Figure 3.1 and additional technical work relevant to each specific data set may be found in Appendix ii. This general approach has also been used for other UKCEH 'Bunce' rescued data sets, which have not yet been described in a peer reviewed journal (Adamson et al., 2017; Bunce et al., 2017a; Bunce et al., 2018a; Bunce et al., 2017c).

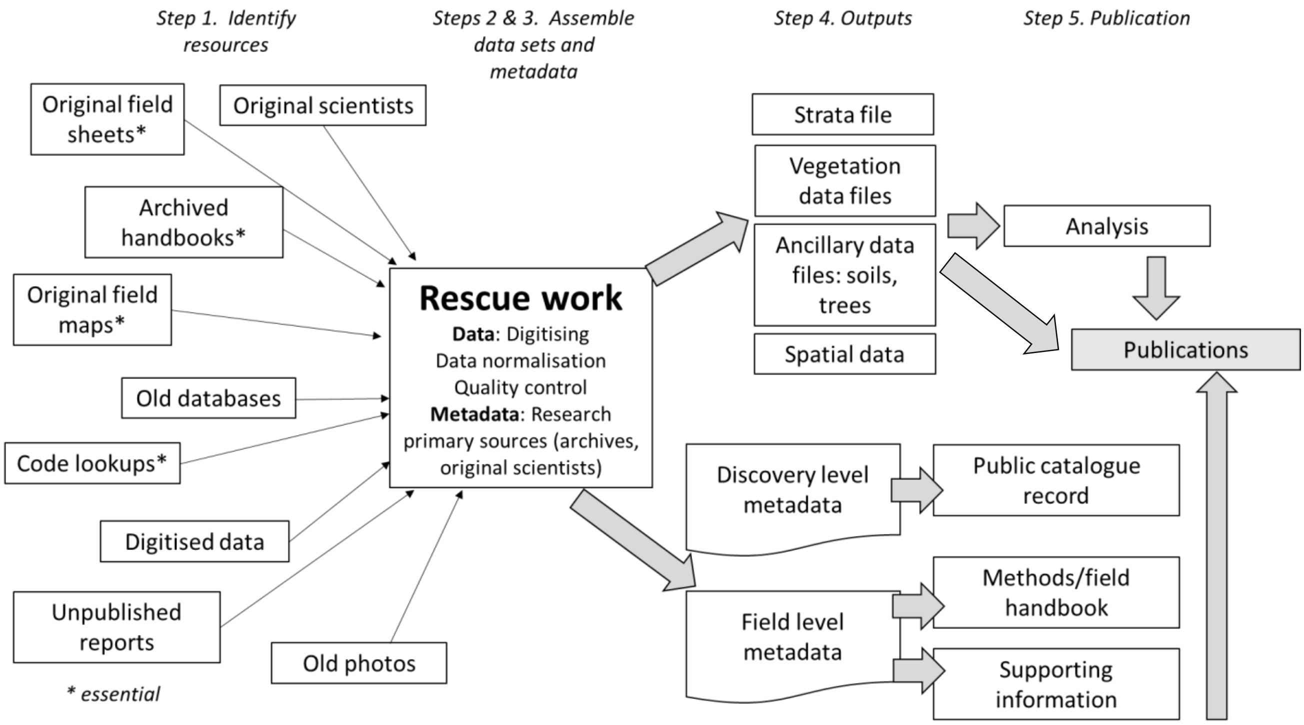


Figure 3.1. Generic approach to the rescue of ecological survey data: workflow

Table 3.1. General approach to rescuing ecological survey data

	Step	***Description***
1	Identify available resources	- Archived field handbooks* - Data on paper recording sheets* - Code lookups* - Paper field maps* - Old photos - Sheets of digitised data in MS Excel spreadsheets - Data in old databases - Unpublished/published reports - Original scientists **essential in order to capture critical information for re-use*
2	Process data sets	- **Plan**: Plan a schema to store data - could the data match a previously designed schema? - **Digitise**: If data entry has not taken place, arrange for entry to be undertaken - Do the site locations need digitising? - **Structure**: Do entered data need re-structuring to allow optimal retrieval and re-use? - **Quality Assurance and Control (QA/QC):** Are there any data missing, based on information given on recording sheets or handbook? Quality checks.
3	Assemble metadata	- **Who:** Is it possible to identify the surveyors from recorded initials? If not, ask originator if possible or check reports/other information.

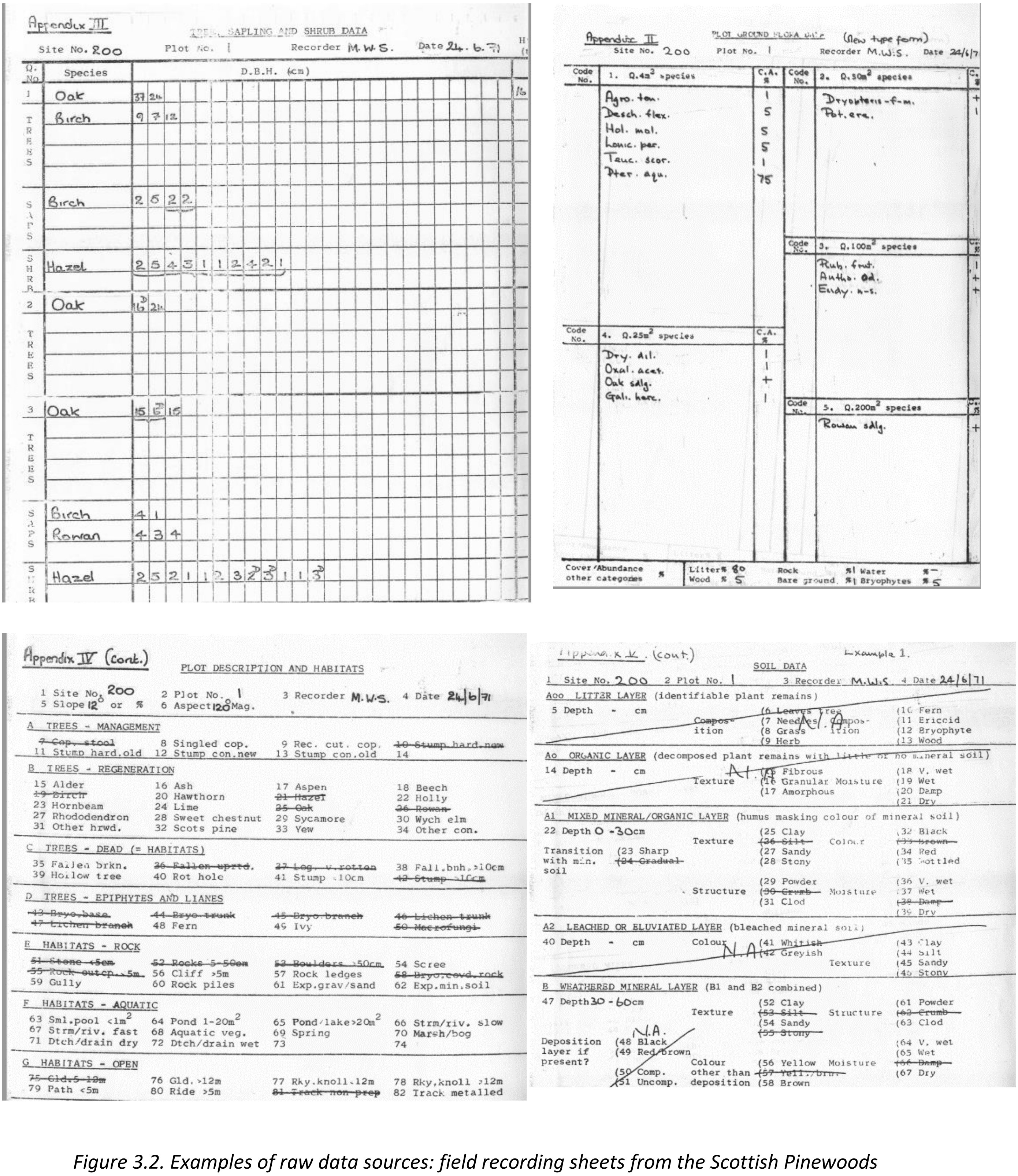
		- **What:** Do the recorded codes have descriptions? Find the code sheets. - Have the correct number of sites/plots been recorded according to the available information? Are there items of data not explained anywhere (e.g. instructions were on an additional sheet). Can any information be found? If not, ask originator if possible - **Where:** Can the site locations be identified? Are there any photographs to help site re-location? - **When:** Are there dates on the recording sheets? Are they consistent for a site? - **How:** Compare the field sheets with the field handbook. Do the recorded data make sense according to what should have been collected? By extension, do the final data sets correspond to the field handbook?
4	Produce outputs	**Outputs:** 1. **Data sets –** *following appropriate formats and filenames.* • The strata file • Vegetation plot data • Ancillary data sampled from plot locations • Mapped feature data 2. **Metadata documents -** *produce metadata documents describing the information discovered in steps 2 and 3. Metadata should follow guidance, as now documented by the NERC Environmental Data Centre (EIDC), and included in Appendix v.* • Discovery metadata (catalogue record) • Field level metadata (supporting documentation)
5	Publication	1. **Published, citable data sets with Digital Object Identifiers (DOIs)** 2. **Data publication in journal** (*optional*) 3. **Analysis outputs** (*optional*)

3.2 Application of the approach, illustrated by the 'Native Pinewoods of Scotland, 1971' data set (Paper 2)

3.2.1 Step 1: Identify available resources

The first stage of rescue is to identify the resources required to aid the rescue. It is of course essential that the data exist in some form. These may be paper recording sheets (Figure 3.2), or the data may have been punched or digitised at some point in the past, perhaps existing in a still accessible spreadsheet format or an old database (Figure 3.3). It is possible that data may have been previously entered into a now obsolete storage medium, in which case, it is often most practical to re-enter the data, assuming the original recording sheets are still in existence. In the case of the UKCEH data, several of the data sets had been punched into MS

Excel spreadsheets in the early 2000s. However they had no metadata, were not structured well, and in that form were essentially useless. In the case of the 'Key Habitats' survey (Paper 6), there were some extant data residing in disparate Oracle databases. These were also extremely badly structured and contained no metadata, and were difficult to find and access. Fortunately, in the case of all the UKCEH data sets, the paper recording sheets were available, which gave crucial clues to the many questions raised from looking at these digitised data sources. In the case of the 'Key Habitats' (Paper 6) and the 1978 Countryside Survey (Paper 5), the data included elements of spatial data containing land use and habitat information which had to be digitised. Without the original sheets, this information might otherwise have been lost.



Appendix III — TREE, SAPLING AND SHRUB DATA

Site No. 200 Plot No. 1 Recorder M.W.S. Date 24.6.71

Q. No		Species	D.B.H. (cm)
1	TREES	Oak	37 24
		Birch	9 7 12
	SAPS	Birch	2 6 2 2
	SHRB	Hazel	2 5 4 3 1 1 2 4 2 1
2	TREES	Oak	16D 24
3	TREES	Oak	15 6D 15
	SAPS	Birch	4 1
		Rowan	4 3 4
	SHRB	Hazel	2 5 2 1 1 2 3 2D 3D 1 1 3D

Appendix II — PLOT GROUND FLORA DATA (New type form)

Site No. 200 Plot No. 1 Recorder M.W.S. Date 24/6/7

Code No.	1. Q.4m² species	C.A. %
	Agro. ten.	1
	Desch. flex.	5
	Hol. mol.	5
	Lonic. per.	5
	Teuc. scor.	1
	Pter. aqu.	75

Code No.	2. Q.50m² species	C.A. %
	Dryopteris-f-m.	+
	Pot. ere.	1

Code No.	3. Q.100m² species	C.A. %
	Rub. frut.	1
	Antho. od.	+
	Endy. n-s.	+

Code No.	4. Q.25m² species	C.A. %
	Dry. dil.	1
	Oxal. acet.	1
	Oak sdlg.	+
	Gali. harc.	1

Code No.	5. Q.200m² species	C.A. %
	Rowan sdlg.	+

Cover/Abundance other categories % — Litter % 80, Wood % 5, Rock %1, Bare ground %1, Water % –, Bryophytes % 5

Appendix IV (cont.) — PLOT DESCRIPTION AND HABITATS

1 Site No. 200 2 Plot No. 1 3 Recorder M.W.S. 4 Date 24/6/71
5 Slope 12° or % 6 Aspect 120° Mag.

A TREES - MANAGEMENT
~~7 Cop. stool~~ 8 Singled cop. 9 Rec. cut. cop. ~~10 Stump hard.new~~ 11 Stump hard.old 12 Stump con.new 13 Stump con.old 14

B TREES - REGENERATION
15 Alder 16 Ash 17 Aspen 18 Beech ~~19 Birch~~ 20 Hawthorn ~~21 Hazel~~ 22 Holly 23 Hornbeam 24 Lime ~~25 Oak~~ ~~26 Rowan~~ 27 Rhododendron 28 Sweet chestnut 29 Sycamore 30 Wych elm 31 Other hrwd. 32 Scots pine 33 Yew 34 Other con.

C TREES - DEAD (= HABITATS)
35 Fallen brkn. ~~36 Fallen uprtd.~~ ~~37 Log. v.rotten~~ 38 Fall.bnh.>10cm 39 Hollow tree 40 Rot hole 41 Stump <10cm ~~42 Stump >10cm~~

D TREES - EPIPHYTES AND LIANES
~~43 Bryo.base~~ ~~44 Bryo.trunk~~ ~~45 Bryo.branch~~ ~~46 Lichen trunk~~ ~~47 Lichen branch~~ 48 Fern 49 Ivy ~~50 Macrofungi~~

E HABITATS - ROCK
~~51 Stone <5cm~~ ~~52 Rocks 5-50cm~~ ~~53 Boulders >50cm~~ 54 Scree ~~55 Rock outcp.<5m~~ 56 Cliff >5m 57 Rock ledges ~~58 Bryo.covd.rock~~ 59 Gully 60 Rock piles 61 Exp.grav/sand 62 Exp.min.soil

F HABITATS - AQUATIC
63 Sml.pool <1m² 64 Pond 1-20m² 65 Pond/lake>20m² 66 Strm/riv. slow 67 Strm/riv. fast 68 Aquatic veg. 69 Spring 70 Marsh/bog 71 Dtch/drain dry 72 Dtch/drain wet 73 74

G HABITATS - OPEN
~~75 Gld.5-12m~~ 76 Gld.>12m 77 Rky.knoll<12m 78 Rky.knoll >12m 79 Path <5m 80 Ride >5m ~~81 Track non-prep~~ 82 Track metalled

Appendix V (cont.) — SOIL DATA — Example 1.

1 Site No. 200 2 Plot No. 1 3 Recorder M.W.S. 4 Date 24/6/71

Aoo LITTER LAYER (identifiable plant remains) — N.A.
5 Depth - cm; Composition (6 Leaves/Tree (7 Needles (8 Grass (9 Herb (10 Fern (11 Ericoid (12 Bryophyte (13 Wood

Ao ORGANIC LAYER (decomposed plant remains with little or no mineral soil) — N.A.
14 Depth - cm; Texture (15 Fibrous (16 Granular (17 Amorphous; Moisture (18 V. wet (19 Wet (20 Damp (21 Dry

A1 MIXED MINERAL/ORGANIC LAYER (humus masking colour of mineral soil)
22 Depth 0 - 30cm; Texture (25 Clay ~~(26 Silt~~ (27 Sandy (28 Stony; Colour (32 Black ~~(33 Brown~~ (34 Red (35 Mottled
Transition with min. soil (23 Sharp ~~(24 Gradual~~
Structure (29 Powder ~~(30 Crumb~~ (31 Clod; Moisture (36 V. wet (37 Wet ~~(38 Damp~~ (39 Dry

A2 LEACHED OR ELUVIATED LAYER (bleached mineral soil) — N.A.
40 Depth - cm; Colour (41 Whitish (42 Greyish; Texture (43 Clay (44 Silt (45 Sandy (46 Stony

B WEATHERED MINERAL LAYER (B1 and B2 combined)
47 Depth 30 - 60cm; Texture (52 Clay ~~(53 Silt~~ (54 Sandy ~~(55 Stony~~; Structure (61 Powder ~~(62 Crumb~~ (63 Clod
Deposition layer if present? (N.A.) (48 Black (49 Red/brown (50 Comp. (51 Uncomp.
Colour other than deposition (56 Yellow ~~(57 Yell./brn.~~ (58 Brown; Moisture (64 V. wet (65 Wet ~~(66 Damp~~ (67 Dry

Figure 3.2. Examples of raw data sources: field recording sheets from the Scottish Pinewoods

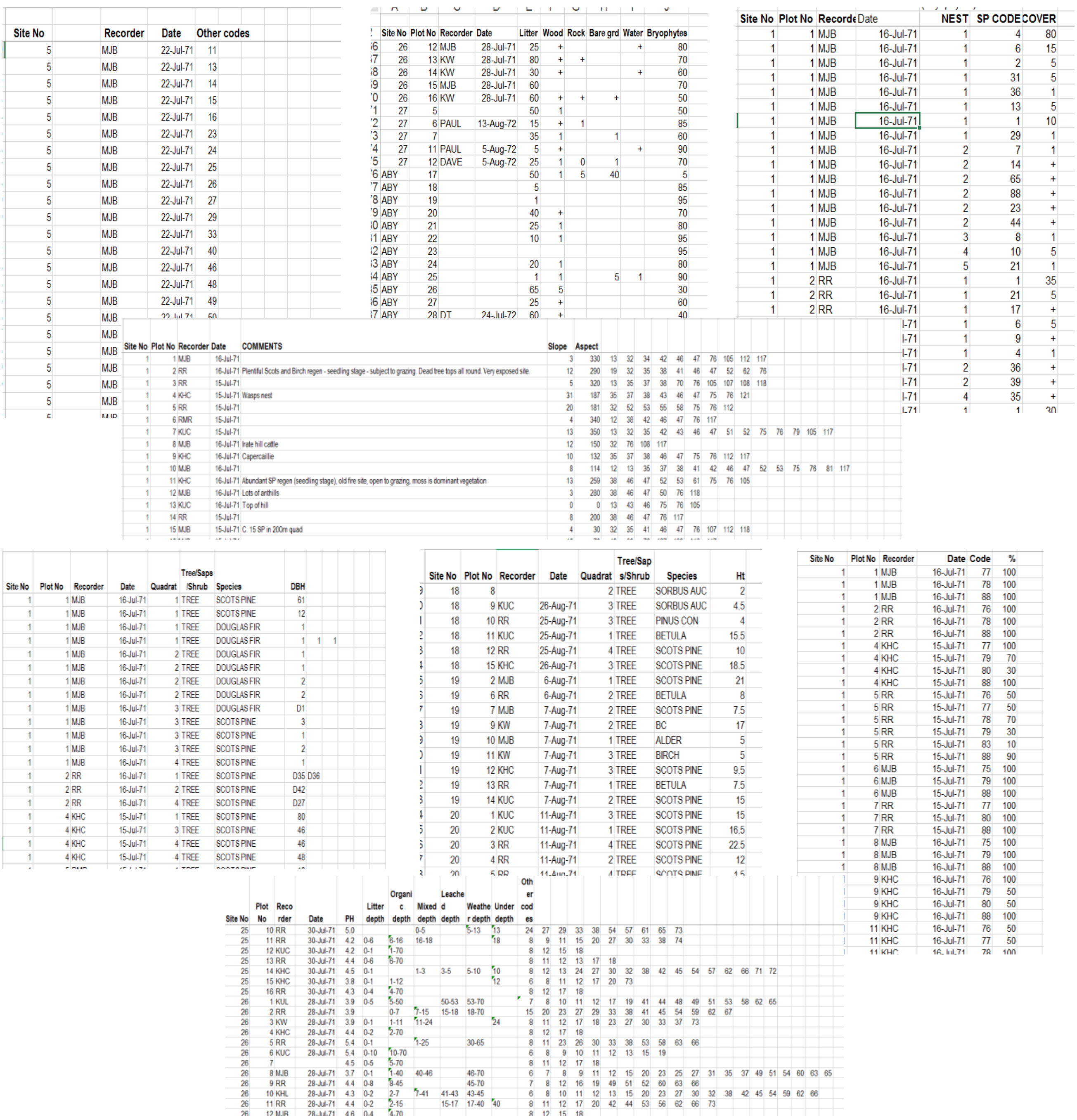


Figure 3.3 Examples of a data source: data previously entered into a spreadsheet

In order that the final metadata are adequate, having the original field handbook is desirable, and quite likely essential, in order to assemble enough information on the data collection methods. A handbook may also contain essential information regarding code descriptions; the data set is rendered useless without this information (Figure 3.4).

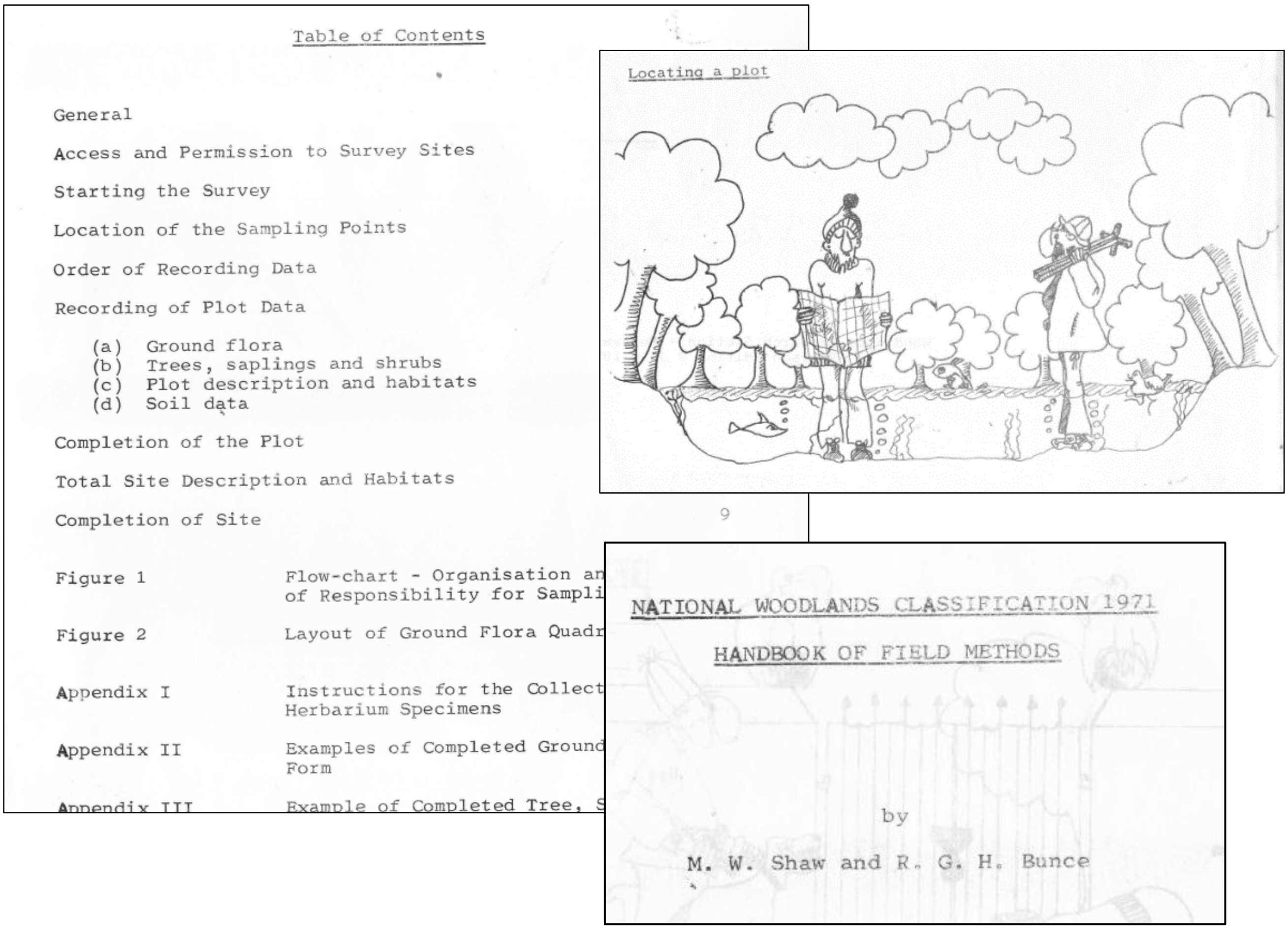


Figure 3.4. A field handbook is an essential source of information

In undertaking a repeat survey, maps of site and plot locations are very valuable. In the case of the UKCEH 'Bunce' surveys, it would be possible to re-randomise new site locations in the event that this information was lost. However, it is of course preferable to be able to relocate the original sites as accurately as possible in order to assess change in an optimal manner because smaller changes can then be identified with a high level of confidence. Photographs and plot sketch maps are also very valuable resources to be able to secure. For a comprehensive set of metadata, it is also extremely useful to be able to consult the originator of the data sets for any details that may be missing from other documentation. Also, a search for unpublished documents or reports can yield useful details (Figure 3.5).

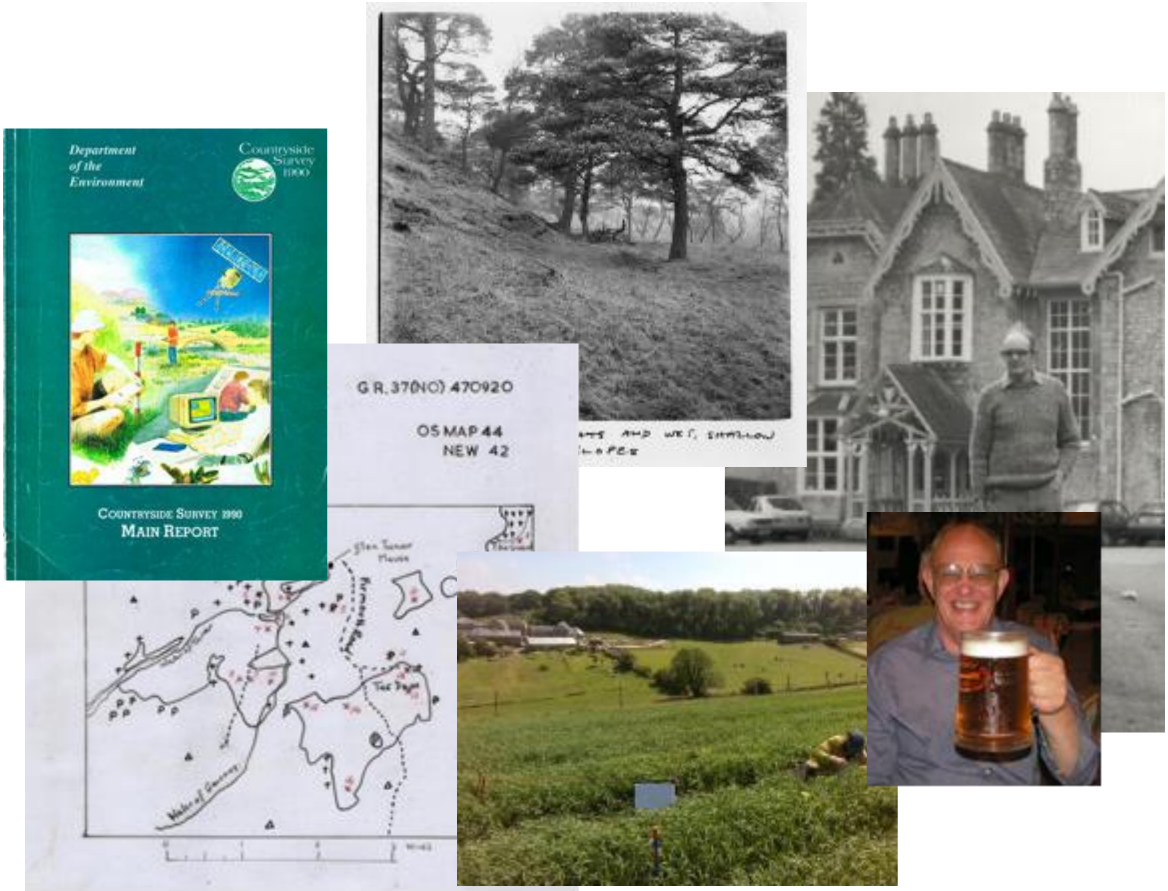


Figure 3.5. Additional sources of information: field photographs, reports, maps, original scientists

3.2.2 Step 2: Process data sets

3.2.2.1 Plan

Steps 2 (Process Data Sets) and 3 (Assemble Metadata) are generally iterative steps, as organising data inevitably raises questions to be explained by the metadata and *vice-versa*. Before commencing any work on the data set, it is important to plan a database schema to load the data into. A database schema refers to the organisation of data within, usually, a relational database and describes a series of database objects such as tables, rows, fields and identifiers, and the relationships between these (Figure 3.6). A formal process of 'database normalisation' ensures minimal data redundancy and improves data integrity (accuracy and consistency) of the data. Database normalisation was first proposed by E.F. Codd in 1970 (Codd, 1970), and consists of a range of rules or processes ('Normal Forms'). Examples of the database normalisation process in relation to the Pinewoods data are given below.

It is also important to consider how users might wish to retrieve data from the database, and then design the schema in such a way as to make retrieval as easy as possible. After designing a schema for one set of data, it is likely to be the case that data collected in a similar manner can be stored in a similar schema. Not only does this make it easier for potential users to become familiar with multiple data sets, it also enhances the possibilities for integrated analyses.

Sometimes, best practice in database design is not always the safest way to proceed in storing ecological legacy data. For example, where database design rules would dictate that the most efficient way of storing data would be to have a numeric code in the main table, and then a lookup table with a key to these codes in a separate table (in order to reduce storage space and data redundancy), it is often safer to store the description in the main table as well. Past experience has shown that code lookup tables may easily be lost or misidentified, hence the meaning of the data in the database is likely to be lost. This is especially true when considering the best way of presenting data in terms of a published product, as opposed to data being stored in internal, institutional relational databases. For publication, it is most efficient to reduce the number of tables presented to end users. It is likely that many end users will not be familiar with interrogating a complicated relational database.

It is also not advisable to enforce referential integrity (whereby records in a related table could be deleted on the basis of their relationship with records in a related table). As ecology is not always an exact science, there are often cases where this might result in a loss of valuable data (for example, as a result of some restriction in the field, a soil sample may have been taken at a plot, but the vegetation may not have been recorded, or *vice-versa*. Enforcing referential integrity might result in losing one or other of these pieces of information).

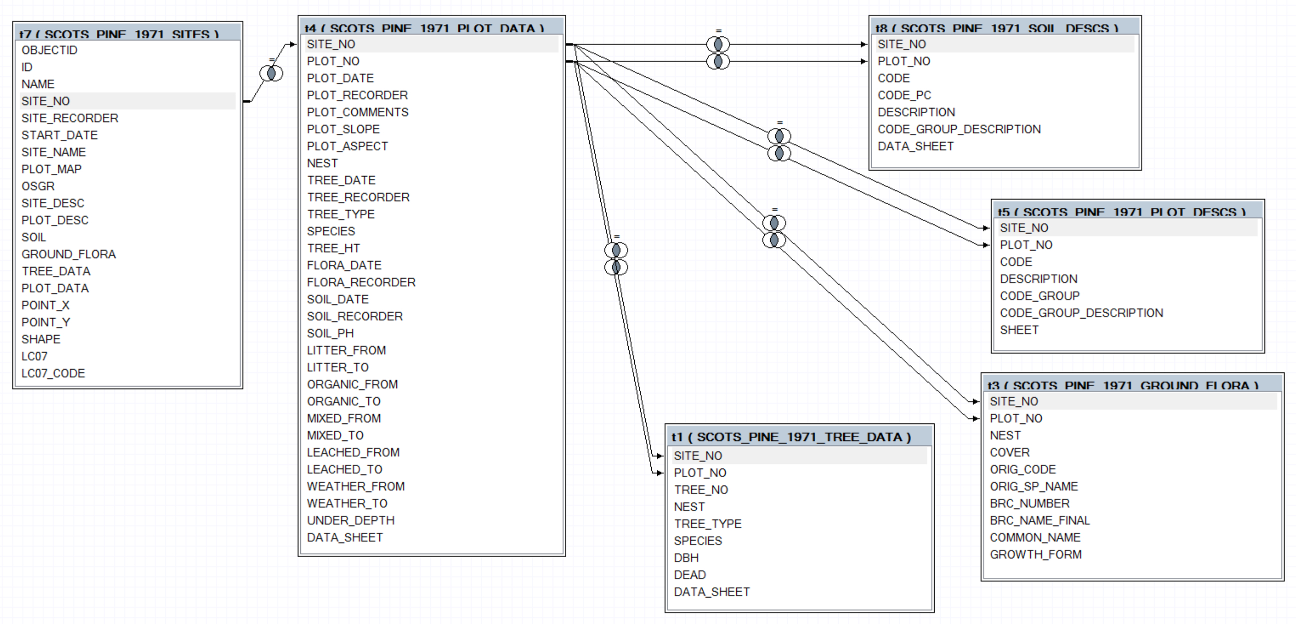


Figure 3.6. Example of a database schema. The Native Pinewoods of Scotland data set

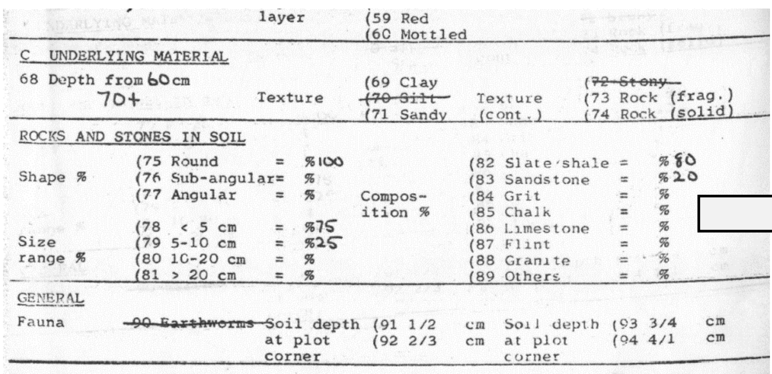
3.2.2.2 Digitise

If the data does not exist at all in a digital form, it is necessary to arrange for entry to be undertaken. This is often the most time-consuming stage of the process if it requires many sheets of field data to be manually typed and checked. In the case of the UKCEH data sets in question, several of them had been typed into spreadsheets in the early 2000s as part of a NERC funded data rescue project (Figure 3.7). However, before data entry, no thought had been given to how the data would be structured or retrieved. In addition, no code descriptions had been stored with the data. Therefore, it took nearly as long to understand and restructure these entered data, as it might have been to start again by entering them in a more structured way.

Despite some of the data already existing in a digital form, key information was still missing. Most importantly, this included the site locations of survey plots. The Shetland Survey, 'Key Habitats' survey and the Countryside Survey also had elements of spatial data to be digitised, as described in Appendix ii).

Whilst NERC had recognised a need for rescuing data back in the 2000s, there are limited options for the funding of this kind of work. Where automated solutions are not possible, useful resources for having data digitised have tended to involve work experience students or interns.

Once the data have been digitised from field sheets, it is also advisable to scan all available documents in order that they can be stored digitally on backed-up network drives, and thus not physically at risk from events such as floods or fire.


layer (59 Red
(60 Mottled
C UNDERLYING MATERIAL
68 Depth from 60 cm 70+ Texture (69 Clay (70 Silt (71 Sandy Texture (cont.) (72 Stony (73 Rock (frag.) (74 Rock (solid)
ROCKS AND STONES IN SOIL
Shape % (75 Round = %100 (76 Sub-angular= % (77 Angular = %
Size range % (78 < 5 cm = %75 (79 5-10 cm = %25 (80 10-20 cm = % (81 > 20 cm = %
Compos-ition % (82 Slate/shale = %80 (83 Sandstone = %20 (84 Grit = % (85 Chalk = % (86 Limestone = % (87 Flint = % (88 Granite = % (89 Others = %
GENERAL
Fauna 90 Earthworms Soil depth at plot corner (91 1/2 cm (92 2/3 cm Soil depth at plot corner (93 3/4 cm (94 4/1 cm

Site No	Plot No	Recorder	Date	Code	%
1	1	MJB	16-Jul-71	77	100
1	1	MJB	16-Jul-71	78	100
1	1	MJB	16-Jul-71	88	100
1	2	RR	16-Jul-71	76	100
1	2	RR	16-Jul-71	78	100
1	2	RR	16-Jul-71	88	100
1	4	KHC	15-Jul-71	77	100
1	4	KHC	15-Jul-71	79	70
1	4	KHC	15-Jul-71	80	30
1	4	KHC	15-Jul-71	88	100
1	5	RR	15-Jul-71	76	50
1	5	RR	15-Jul-71	77	50
1	5	RR	15-Jul-71	78	70
1	5	RR	15-Jul-71	79	30
1	5	RR	15-Jul-71	83	10
1	5	RR	15-Jul-71	88	90
1	6	MJB	15-Jul-71	75	100

Figure 3.7. Digitising data from original field sheets

3.2.2.3 Data structure

As discussed above, in order to ensure optimal re-use and storage of data, a process of structuring and database normalisation needs to be undertaken (see Figure 3.8 for an example). Preferably, this is done at the project planning stage, so data can be digitised into a logical structure. At this point, data entry forms could be specially designed to mimic the storage structure (or be easily manipulated via a scripted process), perhaps with a tool such as Microsoft Access, or ESRI Geographical Information System tools, Survey123 or Collector (ESRI, 2020).

In the case of the Pinewoods survey, data had already been entered into unstructured Microsoft Excel spreadsheets at a point in the past. This meant that in order to organise and manipulate the data into a logical usable structure, a range of normalisation and structuring procedures had to be undertaken. Procedures or rules referred to as 'normal forms' aim to eliminate redundant or useless data, reduce the complexity of the data, ensure the relationship between tables as well as data in the tables, and ensure data dependencies and data are logically stored (SQL World, 2020). Practically, this means eliminating repeating groups in individual tables, creating a separate table for each set of related data and identifying each set of related data with a primary key.

For data to be in the first normal form (1NF), each attribute must be a single value. Therefore only one item of information may be stored in one field or cell. An example of where this was not the case was where tree Diameter at Breast Height data and information relating to whether a tree was dead had been entered and stored within one field (for example '24D', meaning 24cm and D = 'dead'). To satisfy 1NF, these two items of information had to be separated.

The second normal form (2NF) relates to the creation of separate tables for sets of values that apply to multiple records, and relating these tables with a foreign key (a unique identifier in a related table) in order to eliminate redundant data. In the example in Figure 3.8, this means storing the numeric code in the main table and having a separate lookup table (Figure 3.9) (the primary and foreign keys being 'CODE'). In the case of the 'Bunce' surveys, this rule has not entirely been followed due to the potential for confusion between code lists from different surveys, and the potential risk of loss of the lookup table. It was often therefore deemed safer to store the descriptions in the main data tables as well as the code lookup table. When considering how to present the data as a published product, it is also preferable to keep the number of tables to a minimum, from the point of view of an end-user.

The third normal form (3NF) states that fields that do not depend on the key (i.e. SITE and PLOT) should be eliminated. Again, in the example in Figure 3.8, this would dictate that some of the information (e.g. 'CODE_GROUP_DESCRIPTION' and 'DATA_SHEET' should be stored in separate tables, but for the same reason outlined above, this rule has not always been followed. Other normal forms do exist, but are generally not practically implemented (for example, Microsoft (2020)).

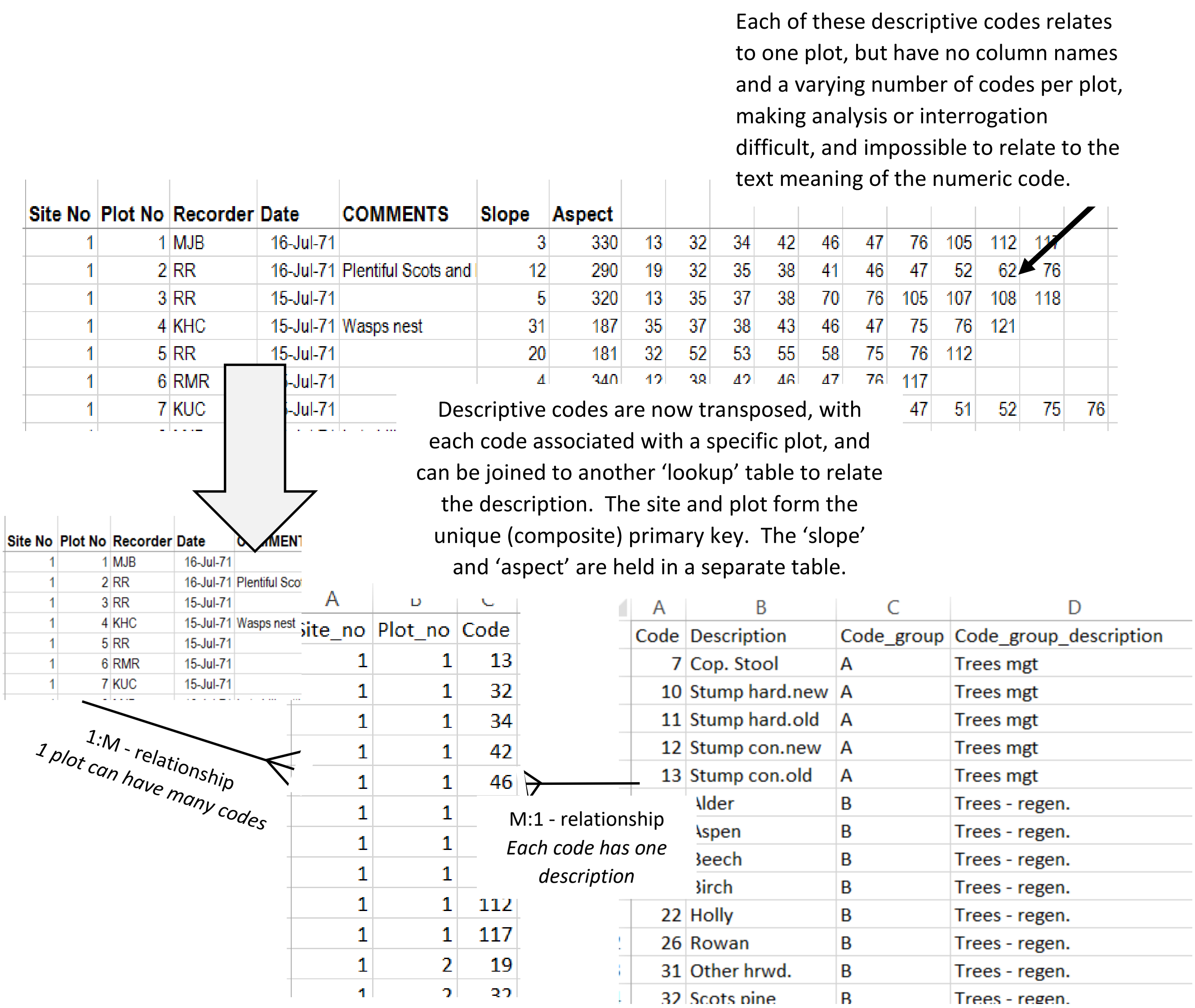
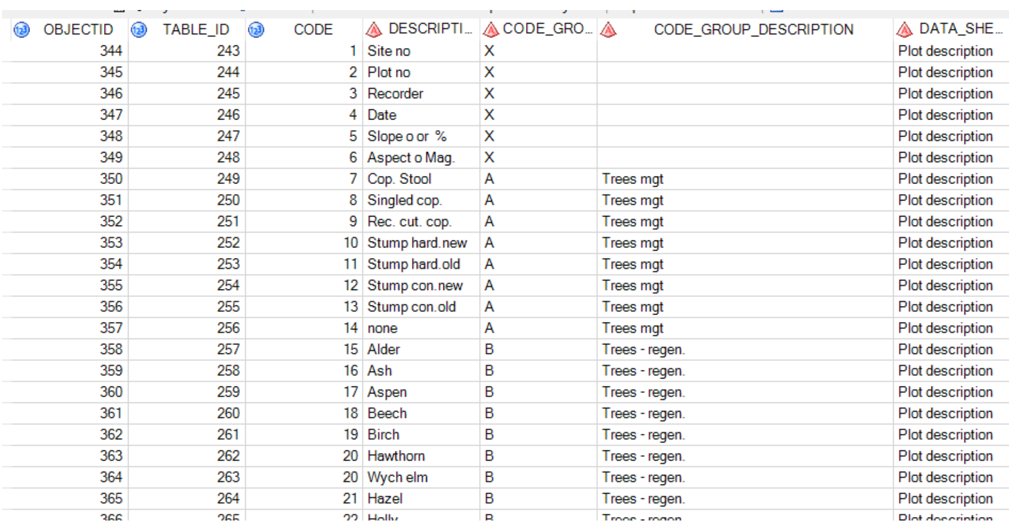


Figure 3.8. Example of transforming unstructured data into a normalised structure.



OBJECTID	TABLE_ID	CODE	DESCRIPTI...	CODE_GRO...	CODE_GROUP_DESCRIPTION	DATA_SHE...
344	243	1	Site no	X		Plot description
345	244	2	Plot no	X		Plot description
346	245	3	Recorder	X		Plot description
347	246	4	Date	X		Plot description
348	247	5	Slope o or %	X		Plot description
349	248	6	Aspect o Mag.	X		Plot description
350	249	7	Cop. Stool	A	Trees mgt	Plot description
351	250	8	Singled cop.	A	Trees mgt	Plot description
352	251	9	Rec. cut. cop.	A	Trees mgt	Plot description
353	252	10	Stump hard.new	A	Trees mgt	Plot description
354	253	11	Stump hard.old	A	Trees mgt	Plot description
355	254	12	Stump con.new	A	Trees mgt	Plot description
356	255	13	Stump con.old	A	Trees mgt	Plot description
357	256	14	none	A	Trees mgt	Plot description
358	257	15	Alder	B	Trees - regen.	Plot description
359	258	16	Ash	B	Trees - regen.	Plot description
360	259	17	Aspen	B	Trees - regen.	Plot description
361	260	18	Beech	B	Trees - regen.	Plot description
362	261	19	Birch	B	Trees - regen.	Plot description
363	262	20	Hawthorn	B	Trees - regen.	Plot description
364	263	20	Wych elm	B	Trees - regen.	Plot description
365	264	21	Hazel	B	Trees - regen.	Plot description
366	265	22	Holly	B	Trees - regen.	Plot description

Figure 3.9. Example of a code 'lookup' table

3.2.2.4 Examples of database normalisation and restructuring procedures from the pinewoods data set

(all of the following examples in Table 3.2 are concerned with structuring the non-normalised data (UNF) data into the normalised form, NF)

Table 3.2. List of example data normalisation procedures, with examples

<table>
<tr><th>Issue</th><th>Example (UNF)</th><th>Correction/Solution (NF)</th><th>Script Ref.</th></tr>
<tr>
<td>Mix of 2 pieces of information in one field (e.g. Tree data – diameter at breast height)</td>
<td>D4 (tree of 4cm Diameter at Breast Height DBH), dead
<table>
<tr><th>Site No</th><th>Plot No</th><th>Recorder</th><th>Date</th><th>Quadrat</th><th>Tree/Saps/Shrub</th><th>Species</th><th>DBH</th><th></th></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>4</td><td>TREE</td><td>SCOTS PINE</td><td>1</td><td></td></tr>
<tr><td>1</td><td>2</td><td>RR</td><td>16-Jul-71</td><td>1</td><td>TREE</td><td>SCOTS PINE</td><td>D35</td><td>D36</td></tr>
<tr><td>1</td><td>2</td><td>RR</td><td>16-Jul-71</td><td>2</td><td>TREE</td><td>SCOTS PINE</td><td>D42</td><td></td></tr>
<tr><td>1</td><td>2</td><td>RR</td><td>16-Jul-71</td><td>4</td><td>TREE</td><td>SCOTS PINE</td><td>D27</td><td></td></tr>
<tr><td>1</td><td>4</td><td>KHC</td><td>15-Jul-71</td><td>1</td><td>TREE</td><td>SCOTS PINE</td><td>80</td><td></td></tr>
<tr><td>1</td><td>4</td><td>KHC</td><td>15-Jul-71</td><td>3</td><td>TREE</td><td>SCOTS PINE</td><td>46</td><td></td></tr>
<tr><td>1</td><td>4</td><td>KHC</td><td>15-Jul-71</td><td>4</td><td>TREE</td><td>SCOTS PINE</td><td>46</td><td></td></tr>
<tr><td>1</td><td>4</td><td>KHC</td><td>15-Jul-71</td><td>4</td><td>TREE</td><td>SCOTS PINE</td><td>48</td><td></td></tr>
<tr><td>1</td><td>5</td><td>RMR</td><td>15-Jul-71</td><td>1</td><td>TREE</td><td>SCOTS PINE</td><td>18</td><td></td></tr>
</table>
</td>
<td>Create 2 separate columns, one with the DBH values and one denoting whether a tree or stem was dead
<table>
<tr><th>SITE_NO</th><th>PLOT_NO</th><th>TREE_NO</th><th>NEST</th><th>TREE_TYPE</th><th>SPECIES</th><th>DBH</th><th>DEAD</th></tr>
<tr><td>1</td><td>1</td><td>8</td><td>2</td><td>TREE</td><td>Pseudotsuga menziesii</td><td>2</td><td></td></tr>
<tr><td>1</td><td>1</td><td>9</td><td>3</td><td>TREE</td><td>Pseudotsuga menziesii</td><td>1</td><td>D</td></tr>
<tr><td>1</td><td>1</td><td>10</td><td>3</td><td>TREE</td><td>Pinus sylvestris</td><td>3</td><td></td></tr>
<tr><td>1</td><td>1</td><td>11</td><td>3</td><td>TREE</td><td>Pinus sylvestris</td><td>1</td><td></td></tr>
<tr><td>1</td><td>1</td><td>12</td><td>3</td><td>TREE</td><td>Pinus sylvestris</td><td>2</td><td></td></tr>
<tr><td>1</td><td>1</td><td>13</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td><td>1</td><td></td></tr>
<tr><td>1</td><td>2</td><td>14</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td><td>36</td><td>D</td></tr>
<tr><td>1</td><td>2</td><td>14</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td><td>35</td><td>D</td></tr>
<tr><td>1</td><td>2</td><td>15</td><td>2</td><td>TREE</td><td>Pinus sylvestris</td><td>42</td><td>D</td></tr>
<tr><td>1</td><td>2</td><td>16</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td><td>27</td><td>D</td></tr>
<tr><td>1</td><td>4</td><td>17</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td><td>80</td><td></td></tr>
<tr><td>1</td><td>4</td><td>18</td><td>3</td><td>TREE</td><td>Pinus sylvestris</td><td>46</td><td></td></tr>
<tr><td>1</td><td>4</td><td>19</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td><td>46</td><td></td></tr>
<tr><td>1</td><td>4</td><td>20</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td><td>48</td><td></td></tr>
<tr><td>1</td><td>5</td><td>21</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td><td>18</td><td></td></tr>
</table>
</td>
<td>SPW1</td>
</tr>
<tr>
<td>Data relationships expressed by brackets (e.g. DBH data, bracketed to denote the 1:M relationship between tree and stems); data cannot be stored in this way.</td>
<td>1222131, 2351


</td>
<td>Brackets denote whether stems were on the same tree. Re-enter data, creating unique tree numbers for each set of data within brackets during data entry.
<table>
<tr><th>SITE_NO</th><th>PLOT_NO</th><th>TREE_NO</th><th>NEST</th><th>TREE_TYPE</th><th>SPECIES</th></tr>
<tr><td>1</td><td>1</td><td>8</td><td>2</td><td>TREE</td><td>Pseudotsuga menziesii</td></tr>
<tr><td>1</td><td>1</td><td>9</td><td>3</td><td>TREE</td><td>Pseudotsuga menziesii</td></tr>
<tr><td>1</td><td>1</td><td>10</td><td>3</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>1</td><td>11</td><td>3</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>1</td><td>12</td><td>3</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>1</td><td>13</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>2</td><td>14</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>2</td><td>14</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>2</td><td>15</td><td>2</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>2</td><td>16</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>4</td><td>17</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>4</td><td>18</td><td>3</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>4</td><td>19</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>4</td><td>20</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>1</td><td>5</td><td>21</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
</table>
</td>
<td>N/A</td>
</tr>
</table>

Additional codes not incorporated into main data set (e.g. Water/bryophytes/litter/rock data, recorded at bottom of ground flora sheets)

Water/bryophytes/litter/rock data, recorded at bottom of ground flora sheets

Site No	Plot No	Recorder	Date	Litter	Wood	Rock	Bare grd	Water	Bryophytes
26	12	MJB	28-Jul-71	25	+			+	80
26	13	KW	28-Jul-71	80	+	+			70
26	14	KW	28-Jul-71	30	+			+	60
26	15	MJB	28-Jul-71	60					70
26	16	KW	28-Jul-71	60	+	+	+		50
27	5			50	1				50
27	6	PAUL	13-Aug-72	15	+	1			85
27	7			35	1		1		60
27	11	PAUL	5-Aug-72	5	+			+	90
27	12	DAVE	5-Aug-72	25	1	0	1		70

Incorporate into ground flora data set, adding to nest 5

SITE_NO	PLOT_NO	NEST	COVER	ORIG_CODE	BRC_NAME_FINAL
1	1	2	0.5	65	Dryopteris dilatata
1	1	2	0.5	88	Rubus chamaemorus
1	1	3	1	8	Agrostis canina sens.lat.
1	1	4	5	10	Pteridium aquilinum
1	1	5	5	.	TOTAL_WOOD
1	1	5	25	.	TOTAL_BRYO
1	1	5	5	.	TOTAL_LITTER

SPW2

Site description codes listed in separate columns in entered data (e.g. Woodland description codes)

13/32/76/85/121/139/181
Stumps – hardwood, old/Tree regeneration/Scree/Small pool/ House occupied/Willow grove/Jay

Site No	Plot No	Recorder	Date	COMMENTS	Slope	Aspect											
1	1	MJB	16-Jul-71		3	330	13	32	34	42	46	47	76	105	112	117	
1	2	RR	16-Jul-71	Plentiful Scots and	12	290	19	32	35	38	41	46	47	52	62	76	
1	3	RR	15-Jul-71		5	320	13	35	37	38	70	76	105	107	108	118	
1	4	KHC	15-Jul-71	Wasps nest	31	187	35	37	38	43	46	47	75	76	121		
1	5	RR	15-Jul-71		20	181	32	52	53	55	58	75	76	112			
1	6	RMR	15-Jul-71		4	340	12	38	42	46	47	76	117				
1	7	KUC	15-Jul-71		13	350	13	32	35	42	43	46	47	51	52	75	76

Transpose codes into one column with associated site/plot codes (and join to code descriptions for the purposes of the end-user)

SITE_NO	PLOT_NO	CODE	DESCRIPTION	CODE_GRO...	CODE_GROUP_DESCRIPTI
2	8	70	Marsh/bog	F	Habitats - aquatic
2	8	76	Gld.>12m	G	Habitats - open
2	8	79	Path <5m	G	Habitats - open
2	8	105	Macfungi.soil	I	Habitats - vegetation
2	8	111	Red deer	J	Animals
2	9	36	Fallen uprtd.	C	Trees - dead
2	9	38	Fall. bnh.>10cm	C	Trees - dead
2	9	46	Lichen trunk	D	Trees - epiphytes and lianes
2	9	47	Lichen branch	D	Trees - epiphytes and lianes
2	9	51	Stone.<5cm	E	Habitats - rock
2	9	52	Rocks 5-50cm	E	Habitats - rock
2	9	53	Boulders >50cm	E	Habitats - rock
2	9	58	Bryo.covd.rock	E	Habitats - rock
2	9	62	Exp.min.soil	E	Habitats - rock
2	9	76	Gld.>12m	G	Habitats - open

SPW3

Mix of 2 pieces of information in one field (soil depths expressed as ranges)

Soil depths expressed as ranges within one field or cell, e.g. Organic depth '4-11'

Split data into an upper and lower value. This then allows calculations of depths and comparisons of horizon boundaries.

N/A

Site No	Plot No	Recorder	Date	PH	Litter depth	Organic depth	Mixed depth	Leached depth	Weather depth	Under depth
1	1	MJB	16-Jul-71	4.1	0-4	4-11	11-70			
1	2	RR	16-Jul-71	4.3	0-1		1-15	15-21	21-38	
1	3	RR	15-Jul-71	4.8	0-2	2-20	20			
1	4	KHC	15-Jul-71	3.9	0-13	13-18	18-25		25-40	
1	5	RR	15-Jul-71	3.7	0-5	5-9	10			
1	6	MJB	15-Jul-71	4.2	0-8	8-13	13-21			21
1	7	RR	15-Jul-71	2.5	0-2	2-10			10	
1	8	MJB	16-Jul-71	3.8	0-4	4-18	18-23		23	
1	9	KHC	16-Jul-71	3.6	0-3	3-18	18-25			25
1	10	RR	16-Jul-71	3.6	0-3	3-17	17-22	22-27	27	
1	11	KHC	16-Jul-71	4.5			0-2	2-12	12-23	23
1	12	MJB	16-Jul-71	3.8	0-4	4-18			18-28	
1	13	KUC	16-Jul-71	3.5	0-4	4-36	36			
1	14	RR	15-Jul-71	3.6	0-3	3-30	30-40		40	
1	15	MJB	15-Jul-71		0-5	5-70				70
1	16	MJB	15-Jul-71	3.9	0-2	2-8	8-13		13-70	70
2	1	MWS	30-Jul-71	5.5	0-2	2-40				40
2	2	MWS	30-Jul-71	3.6	0-4	4-21		21-30		30

LITTER_FROM	LITTER_TO	ORGANIC_FROM	ORGANIC_TO	MIXED_FROM	MIXED_TO
0	1	.	.	1	10
0	2	.	.	14	.
0	1	1	3	3	11
0	1	1	6	6	30
0	5	5	11	.	.
0	1	0	10	10	22
0	2	2	61	.	.
0	.	0	2	2	12
0	.	.	.	.	6
0	1	1	21	.	.
.	.	0	25	.	.
0	1	1	30	.	.
0	1	.	.	1	28
0	10	10	38	.	.
0	1	.	.	1	5
0	4	4	6	6	21
0	4	4	33	.	.
0	10	10	70	.	.

3.2.2.5 Quality assurance/Quality control

Beyond normalisation, additional issues of quality may arise in relation to rescued data sets. Once a data set has been digitised and structured, there are still issues to check, such as: Are there any data missing, based on information given on recording sheets or handbook? Are data stored consistently? Are column names sensible, consistent and free from special characters (e.g.* & £ space ?). Some of these issues are illustrated as below, in Table 3.3.

Table 3.3. List of example data and quality issues encountered, along with solutions

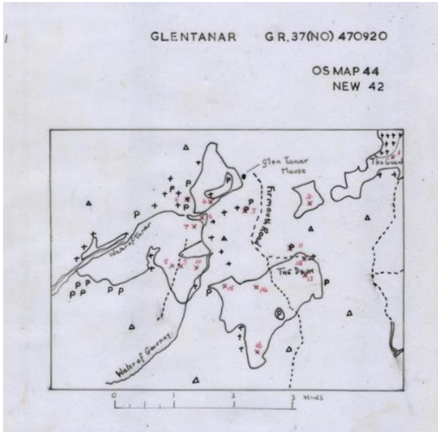
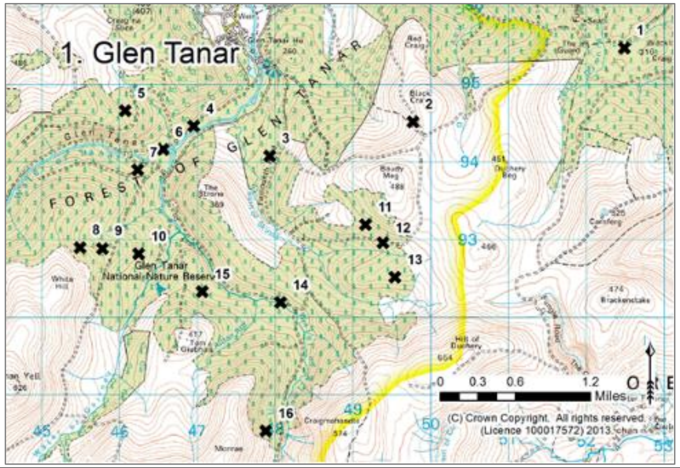
Quality/data issue	Example	Correction/Solution	Script Ref.
Missing codes (e.g. Plot description codes)	Numbers with no entry in the code list	Track down code lists in archive and create new code lookup tables (or defy 2NF and append to data set to be on the safe side – see comments above).	N/A

Site No	Plot No	Recorder	Date	COMMENTS	Slope	Aspect											
1	1	MJB	16-Jul-71		3	330	13	32	34	42	46	47	76	105	112	117	
1	2	RR	16-Jul-71	Plentiful Scots and	12	290	19	32	35	38	41	46	47	52	62	76	
1	3	RR	15-Jul-71		5	320	13	35	37	38	70	76	105	107	108	118	
1	4	KHC	15-Jul-71	Wasps nest	31	187	35	37	38	43	46	47	75	76	121		
1	5	RR	15-Jul-71		20	181	32	52	53	55	58	75	76	112			
1	6	RMR	15-Jul-71		4	340	12	38	42	46	47	76	117				
1	7	KUC	15-Jul-71		13	350	13	32	35	42	43	46	47	51	52	75	76

CODE	DESCRIPTI...	CODE_GRO...	CODE_GROUP_DESCRI...
212	Lake	L	Marginal land use
213	Road	L	Marginal land use
214	Railway	L	Marginal land use
215	Housing	L	Marginal land use
216	Industrial	L	Marginal land use
217	Quarry/mine	L	Marginal land use
218	Tipping	L	Marginal land use
219	Waste	L	Marginal land use
220	None	L	Marginal land use
221	Fence good	M	Boundary type
222	Fence holes	M	Boundary type
223	Fence derelict	M	Boundary type
224	Wall good	M	Boundary type
225	Wall gaps	M	Boundary type

Code lookup table

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr><td></td><td></td><td>
<table>
<tr><th>SITE_NO</th><th>PLOT_NO</th><th>CODE</th><th>DESCRIPTION</th><th>CODE_GRO...</th><th>CODE_GROUP_DESCRIPTI</th></tr>
<tr><td>2</td><td>8</td><td>70</td><td>Marsh/bog</td><td>F</td><td>Habitats - aquatic</td></tr>
<tr><td>2</td><td>8</td><td>76</td><td>Gld.>12m</td><td>G</td><td>Habitats - open</td></tr>
<tr><td>2</td><td>8</td><td>79</td><td>Path <5m</td><td>G</td><td>Habitats - open</td></tr>
<tr><td>2</td><td>8</td><td>105</td><td>Macfungi.soil</td><td>I</td><td>Habitats - vegetation</td></tr>
<tr><td>2</td><td>8</td><td>111</td><td>Red deer</td><td>J</td><td>Animals</td></tr>
<tr><td>2</td><td>9</td><td>36</td><td>Fallen uprtd.</td><td>C</td><td>Trees - dead</td></tr>
<tr><td>2</td><td>9</td><td>38</td><td>Fall. bnh.>10cm</td><td>C</td><td>Trees - dead</td></tr>
<tr><td>2</td><td>9</td><td>46</td><td>Lichen trunk</td><td>D</td><td>Trees - epiphytes and lianes</td></tr>
<tr><td>2</td><td>9</td><td>47</td><td>Lichen branch</td><td>D</td><td>Trees - epiphytes and lianes</td></tr>
<tr><td>2</td><td>9</td><td>51</td><td>Stone.<5cm</td><td>E</td><td>Habitats - rock</td></tr>
<tr><td>2</td><td>9</td><td>52</td><td>Rocks 5-50cm</td><td>E</td><td>Habitats - rock</td></tr>
<tr><td>2</td><td>9</td><td>53</td><td>Boulders >50cm</td><td>E</td><td>Habitats - rock</td></tr>
<tr><td>2</td><td>9</td><td>58</td><td>Bryo.covd.rock</td><td>E</td><td>Habitats - rock</td></tr>
<tr><td>2</td><td>9</td><td>62</td><td>Exp.min.soil</td><td>E</td><td>Habitats - rock</td></tr>
<tr><td>2</td><td>9</td><td>76</td><td>Gld.>12m</td><td>G</td><td>Habitats - open</td></tr>
</table>
Code descriptions appended to data</td><td></td></tr>
<tr><td>Lack of consistent values (e.g. Names of recorded trees)</td><td>Mix of values denoting the same data such as: Betula, Birch, BIRCH, BETULA
<table>
<tr><th>Tree/Saps/Shrub</th><th>Species</th></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
<tr><td>TREE</td><td>BETULA</td></tr>
<tr><td>TREE</td><td>BETULA</td></tr>
<tr><td>TREE</td><td>SILVER BIRCH</td></tr>
<tr><td>TREE</td><td>SILVER BIRCH</td></tr>
<tr><td>TREE</td><td>BETULA</td></tr>
<tr><td>TREE</td><td>BETULA</td></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
<tr><td>TREE</td><td>BIRCH</td></tr>
</table></td><td>Decide on nomenclature (e.g. English or scientific names?) then create lookup table containing all possible values. Use lookup table to replace each set of values with one consistent value.
<table>
<tr><th>SITE_NO</th><th>PLOT_NO</th><th>TREE_NO</th><th>NEST</th><th>TREE_TYPE</th><th>SPECIES</th></tr>
<tr><td>15</td><td>1</td><td>2247</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>1</td><td>2248</td><td>4</td><td>TREE</td><td>Betula sp.</td></tr>
<tr><td>15</td><td>2</td><td>2249</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>2</td><td>2250</td><td>1</td><td>TREE</td><td>Sorbus aucuparia</td></tr>
<tr><td>15</td><td>2</td><td>2251</td><td>2</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>2</td><td>2252</td><td>2</td><td>TREE</td><td>Sorbus aucuparia</td></tr>
<tr><td>15</td><td>2</td><td>2253</td><td>3</td><td>TREE</td><td>Betula sp.</td></tr>
<tr><td>15</td><td>2</td><td>2253</td><td>3</td><td>TREE</td><td>Betula sp.</td></tr>
<tr><td>15</td><td>2</td><td>2254</td><td>3</td><td>TREE</td><td>Betula sp.</td></tr>
<tr><td>15</td><td>2</td><td>2255</td><td>3</td><td>SAPS</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>2</td><td>2256</td><td>3</td><td>SAPS</td><td>Sorbus aucuparia</td></tr>
<tr><td>15</td><td>2</td><td>2257</td><td>3</td><td>SAPS</td><td>Sorbus aucuparia</td></tr>
<tr><td>15</td><td>2</td><td>2258</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>2</td><td>2259</td><td>4</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>2</td><td>2260</td><td>4</td><td>TREE</td><td>Sorbus aucuparia</td></tr>
<tr><td>15</td><td>3</td><td>2261</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>3</td><td>2262</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>3</td><td>2263</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
<tr><td>15</td><td>3</td><td>2264</td><td>1</td><td>TREE</td><td>Pinus sylvestris</td></tr>
</table></td><td>N/A</td></tr>
</table>

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr><td>Mix of data types in one field (e.g. Ground flora, percent cover of species)</td><td>30 (numeric) or + (text)
<table>
<tr><th>Site No</th><th>Plot No</th><th>Recorder</th><th>Date</th><th>NEST</th><th>SP CODE</th><th>COVER</th></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>1</td><td>6</td><td>15</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>1</td><td>13</td><td>5</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>1</td><td>29</td><td>1</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>1</td><td>31</td><td>5</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>1</td><td>36</td><td>1</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>2</td><td>7</td><td>1</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>2</td><td>14</td><td>+</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>2</td><td>23</td><td>+</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>2</td><td>44</td><td>+</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>2</td><td>65</td><td>+</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>2</td><td>88</td><td>+</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>3</td><td>8</td><td>1</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>4</td><td>10</td><td>5</td></tr>
<tr><td>1</td><td>1</td><td>MJB</td><td>16-Jul-71</td><td>5</td><td>21</td><td>1</td></tr>
<tr><td>1</td><td>2</td><td>RR</td><td>16-Jul-71</td><td>1</td><td>1</td><td>35</td></tr>
<tr><td>1</td><td>2</td><td>RR</td><td>16-Jul-71</td><td>1</td><td>4</td><td>1</td></tr>
</table></td><td>Convert + into numeric value (e.g. use 0.5 to denote ‘presence’ of a species). Ensure columns are stored as the correct data type (e.g. a numeric format for percentage values).
<table>
<tr><th>SITE_NO</th><th>PLOT_NO</th><th>NEST</th><th>COVER</th><th>ORIG_CODE</th><th>BRC_NAME_FINAL</th></tr>
<tr><td>1</td><td>1</td><td>1</td><td>10</td><td>1</td><td>Calluna vulgaris</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>5</td><td>2</td><td>Vaccinium myrtillus</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>80</td><td>4</td><td>Deschampsia flexuosa</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>15</td><td>6</td><td>Vaccinium vitis-idaea</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>5</td><td>13</td><td>Galium saxatile</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>1</td><td>29</td><td>Oxalis acetosella</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>5</td><td>31</td><td>Trientalis europaea</td></tr>
<tr><td>1</td><td>1</td><td>1</td><td>1</td><td>36</td><td>Luzula pilosa</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>1</td><td>7</td><td>Potentilla erecta</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>0.5</td><td>14</td><td>Sorbus aucuparia</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>0.5</td><td>23</td><td>Carex binervis</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>0.5</td><td>44</td><td>Holcus lanatus</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>0.5</td><td>65</td><td>Dryopteris dilatata</td></tr>
<tr><td>1</td><td>1</td><td>2</td><td>0.5</td><td>88</td><td>Rubus chamaemorus</td></tr>
</table></td><td>N/A</td></tr>
<tr><td>Location checks</td><td>Plot locations or spatial data missing (may be simple information such as a spot on the map to denote location, or more complex, mapped field data as in the case of the ‘Key Habitat’ survey).

</td><td>Locate field maps, digitise locations, create maps to check sites look in the right place. Store X,Y values in the British National Grid projection format and store in the data table (do not rely on the spatial format of the file).

</td><td>N/A</td></tr>
</table>

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr><td></td><td>(cont'd)</td><td>
<table>
<tr><th>SITE_NO</th><th>PLOT_NO</th><th>POINT_X</th><th>POINT_Y</th></tr>
<tr><td>1</td><td>16</td><td>347877.630402...</td><td>790533.058848...</td></tr>
<tr><td>1</td><td>14</td><td>348069.145679...</td><td>792186.474076...</td></tr>
<tr><td>1</td><td>13</td><td>349537.429473...</td><td>792512.050047...</td></tr>
<tr><td>1</td><td>12</td><td>349384.217251...</td><td>792958.919028...</td></tr>
<tr><td>1</td><td>11</td><td>349160.782760...</td><td>793188.737361...</td></tr>
<tr><td>1</td><td>15</td><td>347060.498552...</td><td>792326.918613...</td></tr>
<tr><td>1</td><td>10</td><td>346243.366702...</td><td>792812.090649...</td></tr>
<tr><td>1</td><td>7</td><td>346230.599017...</td><td>793897.343887...</td></tr>
<tr><td>1</td><td>9</td><td>345777.346193...</td><td>792882.312917...</td></tr>
<tr><td>1</td><td>8</td><td>345490.073277...</td><td>792895.080602...</td></tr>
<tr><td>1</td><td>5</td><td>346071.002952...</td><td>794657.021154...</td></tr>
<tr><td>1</td><td>4</td><td>346951.973228...</td><td>794459.122034...</td></tr>
<tr><td>1</td><td>6</td><td>346556.174988...</td><td>794159.081433...</td></tr>
<tr><td>1</td><td>3</td><td>347928.701143...</td><td>794076.091479...</td></tr>
<tr><td>1</td><td>2</td><td>349767.247805...</td><td>794516.576617...</td></tr>
<tr><td>1</td><td>1</td><td>352486.764744...</td><td>795467.769162...</td></tr>
<tr><td>2</td><td>12</td><td>319490.442951...</td><td>789435.185241...</td></tr>
</table>
</td><td></td></tr>
<tr><td>Column names</td><td>Trees/Saps/Shrubs
<table>
<tr><th>Site No</th><th>Plot No</th><th>Recorder</th><th>Date</th><th>Quadrat</th><th>Tree/Saps/Shrub</th><th>Species</th></tr>
</table>
</td><td>Remove special characters from column names (e.g. /,*,&). Replace spaces with underscores and rename consistently where possible.
<table>
<tr><th>SITE_NO</th><th>PLOT_NO</th><th>TREE_NO</th><th>NEST</th><th>TREE_TYPE</th><th>SPECIES</th></tr>
</table>
</td><td>N/A</td></tr>
</table>

3.2.3 Step 3: Assemble metadata

As discussed in Section 2, it is essential to have an adequate amount of metadata in order to make a data set usable and repeatable. In order to repeat an ecological survey, it is necessary to have 'field level' metadata covering, at a basic level, 'who, what, where, when and how'. EIDC guidance (EIDC, 2020d) also suggests including information on 'why' (Why were the data collected? For what purpose?) and 'completeness' (Are any data absent from the data set? Explain which data are included or excluded and why). The quality of the final data product depends on determining comprehensive answers to these questions.

The following outlines the sort of practical questions to consider whilst undertaking the data rescue tasks (based on the EIDC guidance framework, in bold):

- <u>**Who**</u> - **who collected the data. Were they reliable? Are they still available to ask about discrepancies?**
    - *Is it possible to identify the surveyors from the recorded initials? If not, ask originator if possible or check reports/other information.*
- <u>**What**</u> - **what information was collected exactly (for example definitions, dictionaries and codes)**
    - *Do the recorded codes have descriptions? Find the code lookup sheets.*
    - *Have the correct number of sites/plots been recorded according to the available information?*
- <u>**Where**</u> - **where were the data collected from? (exact grid references).**
    - *Can the site locations be identified from existing maps?*
    - *Are there any photographs in existence to help site re-location? Scan & store securely.*
- <u>**When**</u> - **dates of recording. Seasonal timing is particularly important when considering repeats of botanical data.**
    - *Are there dates on the recording sheets? Are they consistent for each site?*
- <u>**How**</u> **- in what way was the information collected? (handbooks, personal knowledge)**
    - *Compare the field sheets with the field handbook. Do the recorded data make sense according to what should have been collected? By extension, do the final data sets correspond to the field handbook?*
    - *Are there items of data not explained anywhere (for example instructions that were on an additional sheet, and not contained in the handbook). Can any information be found? If not, ask originator if possible.*

Once all possible information has been assembled, it can be inserted into a document, to accompany the published data sets (see Section 3.2.4). As detailed in Section 2, the document should also describe the data in terms of format and physical content, provenance and context, and quality and usefulness.

3.2.4 Step 4: Produce outputs

3.2.4.1 Data

The 'Bunce' data sets, in general, have a similar set of output data sets. These are illustrated in Table 3.4. Most of the data sets have a strata file (or areas of relatively homogenous regions) of the area in question, within which the sample sites are located at random. All of the data sets include ground flora, recorded in a range of plot sizes. All surveys included some sort of soil analysis, except for the 'Key Habitats' – Paper 6. In most cases, this includes laboratory measurements, and often descriptive data recorded in the field. All of the data sets include some kind of assessment of the survey site in terms of land use and habitat descriptions. In the case of the 'Key Habitats' (Paper 6) and the Countryside Survey, this involves more complicated spatial data, whereas the woodlands and Shetland surveys (Papers 1, 2 and 3) included more basic descriptive codes, associated with the plots and sites.

Table 3.4. Summary of 'Bunce' data sets described in the papers

		Woodlands (Paper 1)	**Pinewoods (Paper 2)**	**Shetland (Paper 3)**	**Key Habitats (Paper 6)**	**Countryside Survey (Papers 4,5,7)**
Area of survey in question	**Strata file**	N/A *Later fitted to the ITE Land Classification*	N/A *Full census*	16 strata across Shetland	4 'spatial masks' or strata across England	ITE Land Classification. 45 strata across GB.
Randomly located plots within sites located randomly from strata	**Vegetation plot data**	Vascular plants and some bryophytes	Vascular plants and some bryophytes	Vascular plants and some bryophytes	Vascular plants and some bryophytes	Vascular plants and some bryophytes
Randomly located plots within sites located randomly from strata	**Ancillary data sampled from plot locations**	Soil information, plot descriptions and land use, tree data	Soil information, plot descriptions and land use, tree data	Soil information, plot descriptions	Plot descriptions, land use	Soil information, Plot descriptions
Wider information from site/1km square	**Mapped feature data**	Site descriptions, land use	Site descriptions, land use	Site descriptions	Boundary and area features collected from point locations across 1km survey sites	Boundary, point and area features collected from across 1km square survey sites
	Sampling unit	*Whole woodland site*	*Whole woodland site*	*1km^2*	*1km^2*	*1km^2*

For the internal use of the UKCEH Land Use Group, these data sets are now stored securely in relational tables on the UKCEH Oracle database network in an ArcGIS Enterprise spatial database, and are documented on an internal collaborative ‘wiki’ webpage.

For publication and long-term storage purposes, the data sets have been packaged into downloadable sets of data (Figure 3.10), in non-proprietary formats (in this case, comma separated values files) (which may not follow the structure of the Oracle database exactly, due to a need to incorporate all coded data final tables for the ease of use of end-users, including code descriptions which may be stored in ‘lookup’ tables, as described in Section 3.2.2.3). EIDC offer guidance on this (EIDC, 2020c) and in the case of most tabular data, the preferred format is a comma separated values file. EIDC have also developed guidance on good practice in terms of file structure and naming conventions (EIDC, 2020b). As described in Section 2, the published data sets should be of an acceptable format, should have internally consistent values, should be of sufficient accuracy to enable re-use, should be complete in extent and coverage (according to expectations laid out in the documentation), should contain values that are physically plausible, and should be validated against an independent data set (if possible).

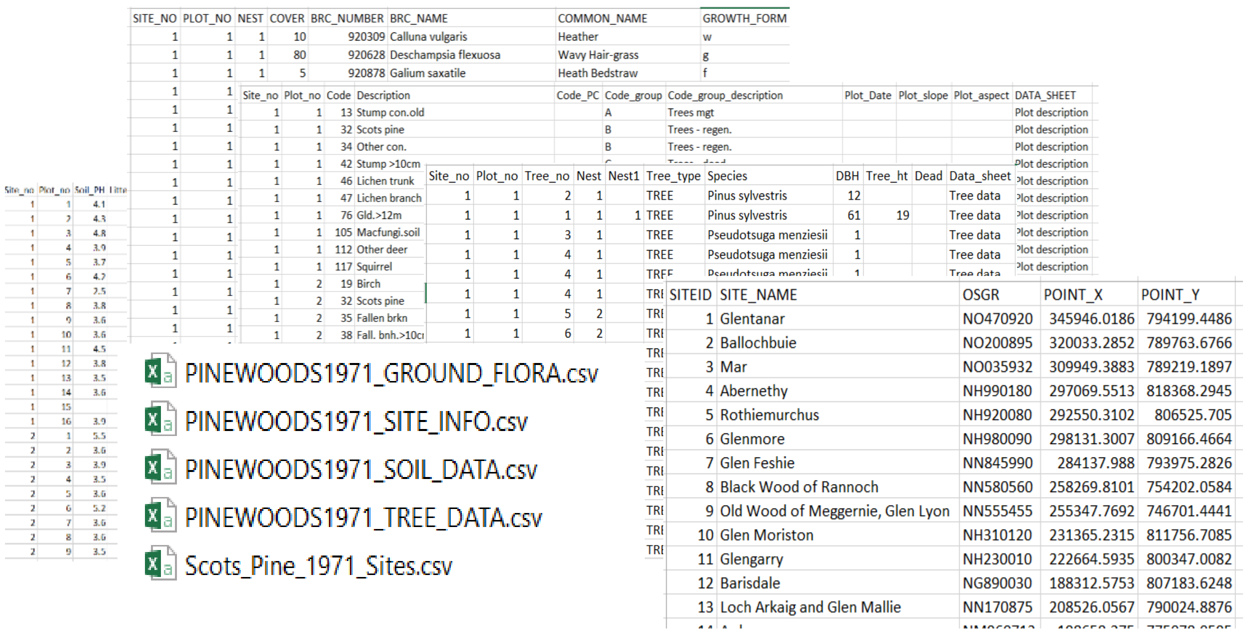


Figure 3.10. Example of the final downloadable data package for the Scottish Pinewoods data set

3.2.4.2 Metadata

The information gathered in Step 3 should be assembled into a user-friendly document, as described in Section 3.2.3. An example of such a document is illustrated in Figure 3.11 and full examples are provided in Appendix iv.

Dataset Documentation

Scottish Pinewoods Survey 1971 (Native Pinewood Survey)

Document version 1.1 4/9/2015
Prepared by C.M. Wood[1], D. Caffrey[2] & R.G.H. Bunce[2].
[1]CEH Lancaster, Library Avenue, Bailrigg, Lancaster. LA1 4AP.
[2] Formerly of the Institute of Terrestrial Ecology, Merlewood, Grange-over-Sands, Cumbria.

Contents

Dataset Series Name:	**Scottish Pinewood Survey 1971**	
Dataset Description:	A detailed ecological survey of the Scots Pine woodland habitats within Scotland. In all, 27 woods from throughout northern Scotland were identified as the major remaining native pinewoods, and within each wood 16 randomly selected 200m^2 plots were surveyed (26 of the woods were surveyed in 1971, with 1 extra wood surveyed in 1972). Details about the trees, ground flora, soil, habitat types as well as general plot information were collected for each plot using standardized procedures and coding systems.	
Geographic Coverage:	Scotland	
Time Period:	1971-72	
Data Categories:	Vegetation Data:	Vascular plants. Bryophytes. Trees, saplings & shrubs.
	Soil Data:	Horizon depths and descriptions. pH.
	Habitat Data:	Habitat categories. Slope. Aspect.
Survey Design & Methods:	Samples of all known major pinewoods in Scotland. Bunce & Shaw's 1973 standardized survey methods.	

Figure 3.11. Example of the supporting metadata document (front page) for the Scottish Pinewoods data set (see Appendix iv for full examples.)

3.2.5 Step 5: Publication (via EIDC https://eidc.ac.uk/)

Once the data and documentation are complete, it becomes straightforward to deposit the package within a long-term data repository, such as the NERC EIDC. The EIDC offer a range of guidance to assist potential depositors, and the basic deposit process is outlined in Table 3.9.

Table 3.5. Guidance from EIDC regarding the data deposit process

How to deposit
We will guide you through the deposit process step-by-step.

Step 1
Contact us
To get started, we recommend you contact us. We will be in touch to find out more about your data such as how it was generated, how was its collection funded and are you working to any deadlines.

Step 2
Agree terms

We will discuss with you what you need to do to prepare, and when you need to do it. This happens before any data is handed over and is formally agreed in a document called a Service Agreement. In it, we agree key details including:

- the format of the data to be deposited
- the volume of data to be handed over
- what documentation is to be provided
- any licensing arrangements
- when the data will be made available to the public (for example, after any embargoes)

Step 3
Get your data and documentation ready
When preparing for deposit, data should be in a non-proprietary format, for example as csvs rather than Excel spreadsheets (see guidance). A full list of file formats we will accept, including advice on the file format you should convert your data to prior to deposit, is available.
You will also be asked to provide appropriate metadata to describe the data and any supporting documentation necessary to its re-use (see guidance).

Step 4
Hand over the data
At the appropriate time, we'll arrange for you to hand over your data and supporting documents.
You can upload your data directly to our website or via Dropbox.com, Google drive, or other file sharing sites
Once you've submitted your data, that's all you need to do! We'll take it from here:
We will check the data and documentation you submit and store it in a secure location which is continuously backed up.
We'll issue a Digital Object Identifier (DOI) for the data, which will enable you and others using your data to cite it (see our DOI policy).
At the appropriate time (subject to any embargo period), we will make your data available via our Data Catalogue. We will also make your data available via other services such as the NERC data service, Google data search, Find open data and the EU INSPIRE portal.

https://eidc.ac.uk/deposit

During this process, EIDC provide a catalogue tool in order to create the 'discovery metadata' introduced in Section 2 (Figure 3.12). Data discovery is the most basic level of metadata. For most studies, a scientist will first be interested in ascertaining whether pertinent data already exist (Michener, 2006). Prior to the 1990s, such data discovery was frequently accomplished via word-of-mouth (for example, presentations at meetings, information exchange with professional colleagues) or the 'methods and materials' section of publications (Michener, 2006Bsarr). As technology has evolved over time, many catalogues now exist. At UKCEH, the EIDC has developed a catalogue over the last decade (https://eip.ceh.ac.uk/data) which has been modified and improved over time. In the case of discovery metadata, a range of standards have been formulated. Within the EU, member states are obliged to curate data sets with metadata conforming at least to the standard of INSPIRE (European Commission, 2020). In the UK, there is the UK GEMINI (GEo-spatial Metadata INteroperability Initiative) (AGI, 2020b) which builds upon the INSPIRE standard. The NERC EIDC adheres to a minimum metadata standard which incorporates the UK GEMINI metadata standard and the NERC metadata quality guidelines. These standards lay out rules for what elements should be recorded regarding the data set (for example: Title, Alternative title, Abstract, Keywords, Temporal extent) and in some cases, what values can be recorded within these elements. Some elements are mandatory, others are optional.

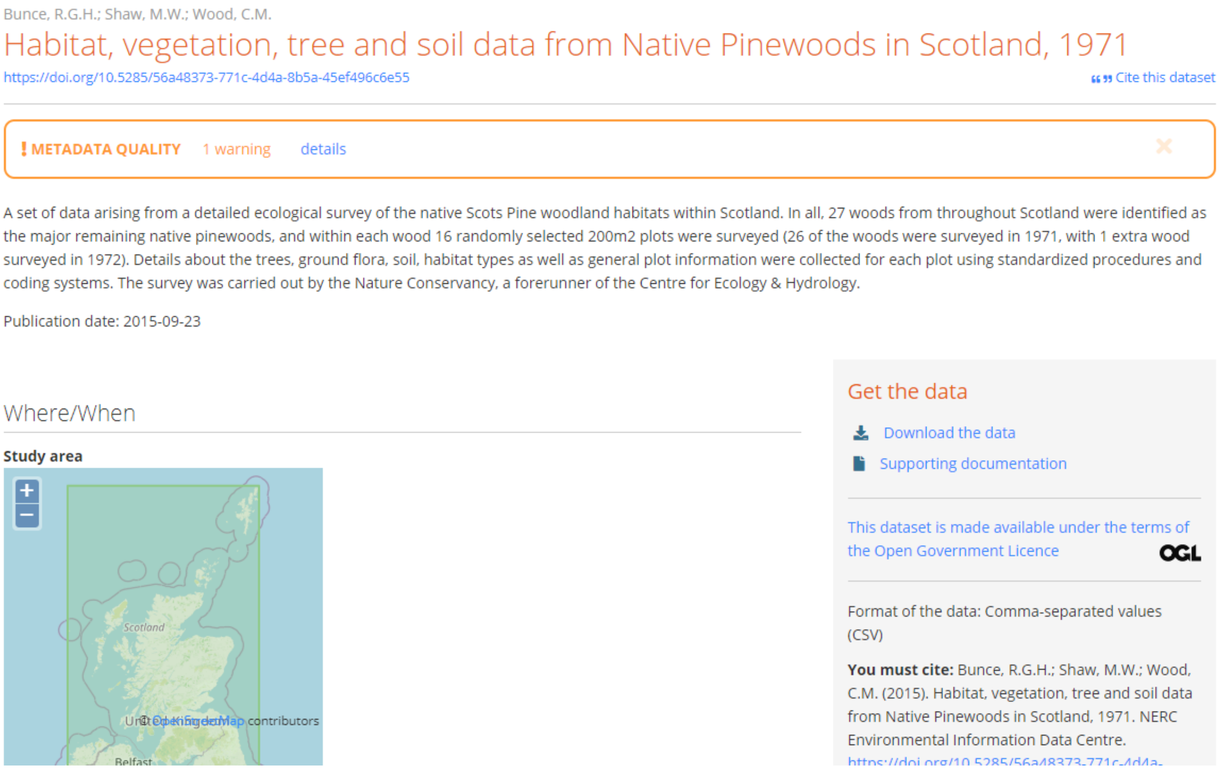


Figure 3.12. Example of an EIDC catalogue record, showing discovery metadata

On successful completion of the data deposit process, deposited data sets are assigned citations, along with unique Digital Object Identifiers (DOIs) which give the data a persistent identifier, which is a guaranteed link to the data for the long-term. The UKCEH 'Bunce' data sets have been assigned DOIs/citations as shown in Table 3.6.

Table 3.6. List of the published data sets with citations described in Papers 1-7

Paper No.	***Survey***	***Data set citations***
1.	Woodland Survey of Great Britain 1971-2001	Kirby, K.J.; Smart, S.M.; Black, H.I.J.; Bunce, R.G.H.; Corney, P.M.; Smithers, R.J.; Shaw, M.W. (2013). **Woodlands survey tree diameter data 1971-2001**. NERC Environmental Information Data Centre. https://doi.org/10.5285/4d93f9ac-68e3-49cf-8a41-4d02a7ead81a Kirby, K.J.; Smart, S.M.; Black, H.I.J.; Bunce, R.G.H.; Corney, P.M.; Smithers, R.J.; Shaw, M.W. (2013). **Woodlands survey site information 1971-2001.** NERC Environmental Information Data Centre. https://doi.org/10.5285/d6409d40-58fe-4fa7-b7c8-71a105b965b4 Kirby, K.J.; Smart, S.M.; Black, H.I.J.; Bunce, R.G.H.; Corney, P.M.; Smithers, R.J.; Shaw, M.W. (2013). **Woodlands survey soil data 1971-2001.** NERC Environmental Information Data Centre. https://doi.org/10.5285/fb1e474d-456b-42a9-9a10-a02c35af10d2 Kirby, K.J.; Smart, S.M.; Black, H.I.J.; Bunce, R.G.H.; Corney, P.M.; Smithers, R.J.; Shaw, M.W. (2013). **Woodlands survey flora data 1971-2001.** NERC Environmental Information Data Centre. https://doi.org/10.5285/2d023ce9-6dbe-4b4f-a0cd-34768e1455ae
2.	Ecological survey of the native pinewoods of Scotland 1971	Bunce, R.G.H.; Shaw, M.W.; Wood, C.M. (2015). **Habitat, vegetation, tree and soil data from Native Pinewoods in Scotland, 1971.** NERC Environmental Information Data Centre. https://doi.org/10.5285/56a48373-771c-4d4a-8b5a-45ef496c6e55
3.	Survey of the terrestrial habitats and vegetation of Shetland, 1974	Bunce, R.G.H.; Bassett, P.A.; Wood, C.M. (2015). **Terrestrial habitat, vegetation and soil data from Shetland, 1974.** NERC Environmental Information Data Centre. https://doi.org/10.5285/06fc0b8c-cc4a-4ea8-b4be-f8bd7ee25342 Bunce, R.G.H.; Bassett, P.A. (2015). **Land Classification of Shetland 1974.** NERC Environmental Information Data Centre. https://doi.org/10.5285/f1b3179e-b446-473d-a5fb-4166668da146
4.	Countryside Survey 1978-2007	Bunce, R.G.H.; Wood, C.M.; Henrys, P.A.; Smart, S.M.; Howard, D.C.; Barr, C.J. (2016). **Landscape area data 1978 [Countryside Survey].** NERC Environmental Information Data Centre. https://doi.org/10.5285/86c017ba-dc62-46f0-ad13-c862bf31740e

5.		Barr, C.J.; Bunce, R.G.H.; Maskell, L.C.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Whittaker, H.A.; Wood, C.M.; Wright, S.M. (2016). **Landscape area data 1984 [Countryside Survey]**. NERC Environmental Information Data Centre. https://doi.org/10.5285/b656bb43-448d-4b2c-aade-7993aa243ea3 Barr, C.J.; Bunce, R.G.H.; Clarke, R.T.; Gillespie, M.K.; Hallam, C.J.; Howard, D.C.; Maskell, L.C.; Ness, M.J.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Wood, C.M.; Wright, S.M. (2016). **Landscape area data 1990 [Countryside Survey]**. NERC Environmental Information Data Centre. https://doi.org/10.5285/94f664e5-10f2-4655-bfe6-44d745f5dca7 Barr, C.J.; Bunce, R.G.H.; Clarke, R.T.; Gillespie, M.K.; Howard, D.C.; Maskell, L.C.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Watkins, J.W.; Wood, C.M.; Wright, S.M. (2016). **Landscape area data 1998 [Countryside Survey]**. NERC Environmental Information Data Centre. https://doi.org/10.5285/1e050028-5c55-42f4-a0ea-c895d827b824 Brown, M.J.; Bunce, R.G.H.; Carey, P.D.; Chandler, K.; Crowe, A.; Maskell, L.C.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Wood, C.M.; Wright, S.M. (2016). **Landscape area data 2007 [Countryside Survey].** NERC Environmental Information Data Centre. https://doi.org/10.5285/bf189c57-61eb-4339-a7b3-d2e81fdde28d
7.		Barr, C.J.; Bunce, R.G.H.; Maskell, L.C.; Norton, L.R.,; Scott, R.J.; Scott, W.A.; Smart, S. M.; Stuart, R.C.; Whittaker, H.A.; Wood, C.M.; Wright, S.M. (2016**). Landscape linear feature data 1984 [Countryside Survey].** NERC Environmental Information Data Centre. https://doi.org/10.5285/a3f5665c-94b2-4c46-909e-a98be97857e5 Barr, C.J.; Bunce, R.G.H.; Clarke, R.T.; Gillespie, M.K.; Hallam, C.J.; Howard, D.C.; Maskell, L.C.; Ness, M.J.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Wood, C.M.; Wright, S.M. (2016). **Landscape linear feature data 1990 [Countryside Survey]**. NERC Environmental Information Data Centre. https://doi.org/10.5285/311daad4-bc8c-485a-bc8a-e0d054889219 Barr, C.J.; Bunce, R.G.H.; Clarke, R.T.; Gillespie, M.K.; Howard, D.C.; Maskell, L.C.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Watkins, J.W. ; Wood, C.M.; Wright, S.M. (2016). **Landscape linear feature data 1998 [Countryside Survey].** NERC Environmental Information Data Centre. https://doi.org/10.5285/8aaf6f8c-c245-46bb-8a2a-f0db012b2643

		Brown, M.J.; Bunce, R G.H.; Carey, P.D.; Chandler, K.; Crowe, A.; Maskell, L.C.; Norton, L.R., ; Scott, R.J.; Scott, W.A.; Smart, S. M.; Stuart, R.C.; Wood, C.M.; Wright, S.M. (2016). **Landscape linear feature data 2007 [Countryside Survey].** NERC Environmental Information Data Centre. https://doi.org/10.5285/e1d31245-4c0a-4dee-b36c-b23f1a697f88 Barr, C.J.; Bunce, R.G.H.; Maskell, L.C.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Whittaker, H.A.; Wood, C.M.; Wright, S.M. (2016). **Landscape point feature data 1984 [Countryside Survey]**. NERC Environmental Information Data Centre. https://doi.org/10.5285/124b872e-036e-4dd3-8316-476b5f42c16e Barr, C.J.; Bunce, R.G.H.; Clarke, R.T.; Gillespie, M.K.; Hallam, C.J.; Howard, D.C.; Maskell, L.C.; Ness, M.J.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Wood, C.M.; Wright, S.M. (2016). **Landscape point feature data 1990 [Countryside Survey].** NERC Environmental Information Data Centre. https://doi.org/10.5285/1481bc63-80d7-4d18-bcba-8804aa0a9e1b Barr, C.J.; Bunce, R.G.H.; Clarke, R.T.; Gillespie, M.K.; Howard, D.C.; Maskell, L.C.; Norton, L.R.; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Watkins, J.W. ; Wood, C.M.; Wright, S.M. (2016). **Landscape point feature data 1998 [Countryside Survey].** NERC Environmental Information Data Centre. https://doi.org/10.5285/ed10944f-40c8-4913-b3f5-13c8e844e153 Brown, M.J.; Bunce, R.G.H.; Carey, P.D.; Chamberlain, P.M.; Chandler, K.; Crowe, A.; Maskell, L.C.; Norton, L.R., ; Scott, R.J.; Scott, W.A.; Smart, S.M.; Stuart, R.C.; Wood, C.M.; Wright, S.M. (2016). **Landscape point feature data 2007 [Countryside Survey]**. NERC Environmental Information Data Centre. https://doi.org/10.5285/55dc5fd7-d3f7-4440-b8a7-7187f8b0550b Barr, C.J.; Bunce, R.G.H.; Smart, S.M.; Whittaker, H.A. (2014). **Countryside Survey 1978 vegetation plot data.** NERC Environmental Information Data Centre. https://doi.org/10.5285/67bbfabb-d981-4ced-b7e7-225205de9c96 Barr, C.J.; Bunce, R.G.H.; Gillespie, M.K.; Hallam, C.J.; Howard, D.C.; Maskell, L.C.; Ness, M.J.; Norton, L.R.; Scott, R.J.; Smart, S.M.; Stuart, R.C.; Wood, C.M. (2014). **Countryside Survey 1990 vegetation plot data.** NERC Environmental Information Data Centre. https://doi.org/10.5285/26e79792-5ffc-4116-9ac7-72193dd7f191 Barr, C.J.; Bunce, R.G.H.; Gillespie, M.K.; Howard, D.C.; Maskell, L.C.; Norton, L.R.; Scott, R.J.; Shield, E.R.; Smart, S.M.; Stuart, R.C.;

		Watkins, J.W. ; Wood, C.M. (2014**). Countryside Survey 1998 vegetation plot data.** NERC Environmental Information Data Centre. https://doi.org/10.5285/07896bb2-7078-468c-b56d-fb8b41d47065 Bunce, R.G.H.; Carey, P.D.; Maskell, L.C.; Norton, L.R.; Scott, R.J.; Smart, S.M.; Wood, C.M. (2014). **Countryside Survey 2007 vegetation plot data.** NERC Environmental Information Data Centre. https://doi.org/10.5285/57f97915-8ff1-473b-8c77-2564cbd747bc Bunce, R.G.H.; Barr, C.J.; Clarke, R.T.; Howard, D.C.; Scott, W.A. (2007**). ITE Land Classification of Great Britain 2007.** NERC Environmental Information Data Centre. https://doi.org/10.5285/5f0605e4-aa2a-48ab-b47c-bf5510823e8f
6.	Ecological survey of \`Key Habitat' landscapes in England, 1992-1993	Barr, C.J.; Bunce, R.G.H.; Cummins, R.P.; Hallam, C.J.; Hornung, M.; Wood, C.M. (2017). **Habitat and vegetation data from an ecological survey of terrestrial Key Habitats in England, 1992-1993.** NERC Environmental Information Data Centre. https://doi.org/10.5285/7aefe6aa-0760-4b6d-9473-fad8b960abd4 Bunce, R.G.H.; Parr, T.W.; Ullyett, J.; Hornung, M.; Gerard, F.; Bull, R.; Cox, R.; Brown, N.J. (2017). **Spatial masks for calcareous, coastal, upland and lowland heath landscapes in England [Key Habitats 1992-93].** NERC Environmental Information Data Centre. https://doi.org/10.5285/dc583be3-3649-4df6-b67e-b0f40b4ec895

Note: See Appendix ii for more details of my contribution to the rescue of each data set.

The citable data sets can then be referenced in any publication as required, including scientific and data journal outputs. In publishing in a data journal, the data sets are fully peer-reviewed and at that point, essentially deemed fit for re-use.

3.2.6 Analysis outputs (optional step)

Once satisfactory data rescue has been completed by following the steps above, the newly available data can be used to perform analyses, as required. Initial analysis may be undertaken at the data quality check stage. For example, a simple distribution graph may help to identify outliers in data which may lead to a need to double check an entered value.

Within Paper 2, regarding the Scottish Pinewoods, a simple analysis of the 25 most frequent vascular plants was undertaken using the ground flora data (Script SPW4, Table 3.7).

Table 3.7. The 25 most frequent vascular plants in the Pinewoods

	Species	Common name	No. of records
1	*Calluna vulgaris*	Heather	398
2	*Vaccinium myrtillus*	Bilberry	356
3	*Potentilla erecta*	Tormentil	333
4	*Deschampsia flexuosa*	Wavy hair grass	299
5	*Vaccinium vitis-idaea*	Cowberry	287
6	*Molinia caerulea*	Purple moor-grass	286
7	*Blechnum spicant*	Hard fern	272
8	*Betula* sp.	Birch	233
9	*Agrostis canina*	Brown bent	214
10	*Pteridium aquilinum*	Bracken	212
11	*Erica tetralix*	Cross-leaved heather	210
12	*Carex echinata*	Star sedge	195
13	*Sorbus aucuparia*	Rowan	194
14	*Galium saxatile*	Heath bedstraw	186
15	*Narthecium ossifragum*	Bog asphodel	169
16	*Luzula multiflora*	Heath woodrush	156
17	*Erica cinerea*	Bell heather	145
18	*Pinus sylvestris*	Scots pine	142
19	*Viola riviniana/reichenbachiana*	Common dog violet	137
20	*Melampyrum pratense*	Common cow-wheat	130
21	*Polygala serpyllifolia*	Heath milkwort	127
22	*Succisa pratensis*	Devil's bit scabious	127
23	*Carex panicea*	Carnation sedge	115
24	*Carex binervis*	Green-ribbed sedge	113
25	*Eriophorum vaginatum*	Hare's-tail cottongrass	111

A set of frequency diagrams illustrating tree diameter within the different woods was also undertaken, an example of which is shown in Figure 3.13 (Script SPW5).

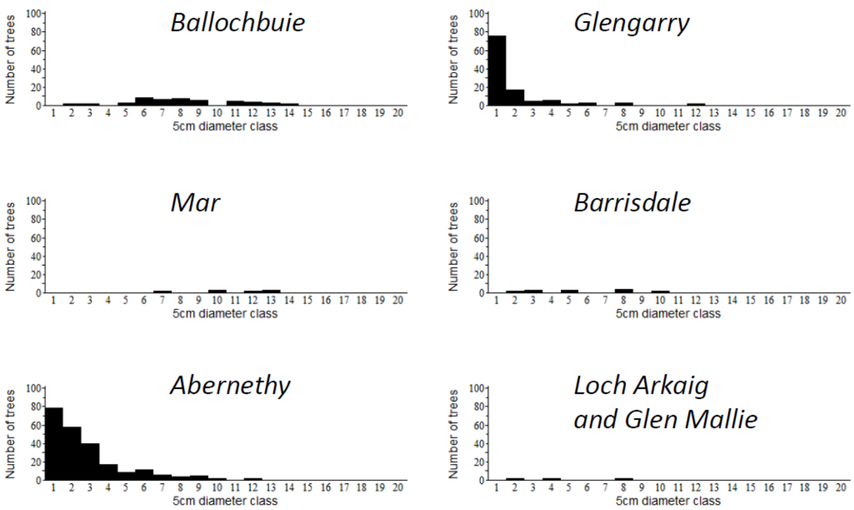


Figure 3.13. Tree diameter distribution of 6 of the Scottish Pinewoods

A range of additional examples of analyses undertaken in relation to the other 'Bunce' data sets are presented in Appendix ii.

4. Publications Summary

This section contains a summary of my submitted papers, outlining their wider significance, my contribution to the work and the links between them. An overview is given in Table 4.1.

Table 4.1 An overview of how the different surveys and data sets fit into the overall timeframe, 1971 – present

Paper No.	Survey	Year begun	Repeated	Significance of survey	Citation
1.	Woodland Survey of Great Britain 1971-2001	**1971**	**2000-03, 2020-22**	*First time the basic methodology was used in earnest, within a single land cover type (following a small pilot woodland survey in the Lake District in 1969, Bunce et al (2018). Repeated 2000-2003 (broadleaved woodlands only), 2020-2023 (both).*	Wood, C.M.; Smart, S.M.; Bunce, R.G.H. (2015) **Woodland Survey of Great Britain 1971-2001.** Earth System Science Data, 7 (2). 203-214. https://doi.org/10.5194/essd-7-203-2015
2.	Ecological survey of the native pinewoods of Scotland 1971	**1971**	**2020-22**		Wood, Claire M.; Bunce, Robert G.H. (2016) **Ecological survey of the native pinewoods of Scotland 1971**. Earth System Science Data, 8 (1). 177-189. https://doi.org/10.5194/essd-8-177-2016
3.	Survey of the terrestrial habitats and vegetation of Shetland, 1974	**1973**	**Not yet**	*First time the stratified random sampling strategy was used across the full range of land cover types (followed by a similar survey in Cumbria (Bunce and Smith, 1978)).*	Wood, Claire M.; Bunce, Robert G.H. (2016) **Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring.** Earth System Science Data, 8 (1). 89-103. https://doi.org/10.5194/essd-8-89-2016
4.	Countryside Survey 1978-2007	**1978**	**1984, 1990, 1998, 2007, 2019-**	*First time the stratified random sampling strategy was used for all land cover types across a whole country. Repeated 1984, 1990, 1998, 2007, 2019-*	Wood, Claire M.; Bunce, Robert G.H.; Norton, Lisa R.; Maskell, Lindsay C.; Smart, Simon M.; Scott, W. Andrew; Henrys, Peter A.; Howard, David C.; Wright, Simon M.; Brown, Michael J.; Scott, Rod J.; Stuart, Rick C.; Watkins, John W. (2018) **Ecological landscape elements: long-term monitoring in Great Britain, the Countryside Survey 1978-2007 and beyond.** Earth System Science Data, 10 (2). 745-763.

					https://doi.org/10.5194/essd-10-745-2018
5.					Wood, Claire M.; Smart, Simon M.; Bunce, Robert G.H.; Norton, Lisa R.; Maskell, Lindsay C.; Howard, David C.; Scott, W. Andrew; Henrys, Peter A. (2017) **Long-term vegetation monitoring in Great Britain - the Countryside Survey 1978-2007 and beyond**. Earth System Science Data, 9 (2). 445-459. https://doi.org/10.5194/essd-9-445-2017
7.					Bunce, Robert G.H.; Wood, Claire M.; Smart, Simon M. (2018) **The ecology of British upland landscapes. I. Composition of landscapes, habitats, vegetation and species.** Journal of Landscape Ecology, 11 (3). 120-139. https://doi.org/10.2478/jlecol-2018-0015
6.	Ecological survey of `key habitat' landscapes in England, 1992-1993	**1992**	**Not yet**	*Example of how the basic methodologies can be adapted to allow a focus on rarer habitat and landscape types.*	Wood, Claire M.; Bunce, Robert G.H.; Norton, Lisa R.; Smart, Simon M.; Barr, Colin J. (2018) **Land cover and vegetation data from an ecological survey of `key habitat' landscapes in England, 1992-1993.** Earth System Science Data, 10 (2). 899-918. https://doi.org/10.5194/essd-10-899-2018

1. Wood, C.M.; Smart, S.M.; Bunce, R.G.H. (2015) **Woodland Survey of Great Britain 1971-2001.** Earth System Science Data, 7 (2). 203-214. https://doi.org/10.5194/essd-7-203-2015

Summary

This paper concerns the Woodland Survey of Great Britain, a unique data set, consisting of a detailed range of ecological measurements at a national scale, covering a time span of 30 years. A set of 103 woods spread across Britain were first surveyed in 1971, which were again surveyed in 2000–2003. Standardised methods of describing the trees, shrubs, ground flora, soils and general habitats present were used for both sets of surveys. The sample of 1648 plots spread through 103 woodland sites located across Britain makes it the most extensive quantitative ecological woodland survey undertaken in Britain; it is also notable for the range of sites that have been revisited after such a long interval.

Contribution to field

There are no other data sets available which provide integrated ground flora, soils and tree data from a range of British woodlands on such a timescale, with the potential to be repeated to investigate change. Having this data well documented and available provides a unique opportunity to repeat the survey for a third time, in order to analyse long-term changes in this habitat at a national scale. Before the publication of this paper, the data, methods and descriptions were not publicly available or described in such a way to enable the re-use of the information.

Links to my other works

This paper has close links to the survey of Native Pinewoods (Wood & Bunce, 2016a). Both of these surveys were undertaken at the same time, using the same methodology. Additionally, the methods developed in the woodland survey were used to create a framework for subsequent surveys (Papers 2-7, Adamson et al. (2017); Bunce et al. (2017a); Emmett and GMEP team (2017)). In terms of the data rescue, this was the first attempt to synthesise and publish the range of information necessary to describe a survey in full, including the methods, data descriptions and quality. It was used as a blueprint for the following five papers.

Work undertaken

As the survey had been repeated in 2001, the data were not at risk as much as most of the other examples presented here. Whilst Bunce (1982) and Kirby et al. (2005) undertook the original surveys, my task was to present the data in such a way as to render it re-usable, collate and research the methods used for data collection, understand the wider implications of the survey and synthesise the information for re-use. The data sets were among the first published by the EIDC and contributed to testing and developing the procedures put in place to do this.

Journal standing and citations

Earth System Science Data (ESSD) is an international, interdisciplinary journal for the publication of articles on original research data (sets), furthering the re-use of high-quality data of benefit to Earth system science. It has an impact factor of 10.95.

To date this article has eight citations.

2. Wood, Claire M.; Bunce, Robert G.H. (2016) **Ecological survey of the native pinewoods of Scotland 1971.** Earth System Science Data, 8 (1). 177-189. https://doi.org/10.5194/essd-8-177-2016

Summary

This paper concerns the 1971 survey of 27 Scottish Pinewoods undertaken across Scotland in 1971. The survey was initiated as a consequence of growing concern about the status of the pinewood resource. Since the twentieth century, this unique habitat has been widely recognised, not only by ecologists for its inherent biodiversity but also by the general public for its cultural and amenity value. This survey varies slightly from the other surveys in question, in that nearly the entire population of the habitat was sampled, rather than using the stratification/environmental classification method, being the major 27 sites of the 35 sites identified as truly native pinewoods in Scotland by Steven and Carlisle (1959). However, the survey utilised the same repeatable methods as used in the Woodland survey, collecting information on ground flora, soils, forest structure and also general site information from each forest. The results from the survey prompted the organisation of an international symposium in 1975, which set the conservation agenda for the old Caledonian pinewoods. Although the data are now 49 years old, the repeatable methods will allow a resurvey to take place, in order to assess changes in the vegetation, habitats and tree composition in a statistically robust manner.

Contribution to field

There are no other data sets available, which provide ground flora, soils and tree data from the Scottish Pinewoods on such a timescale. Having this data set well documented and available provides an excellent opportunity to repeat the survey and provide an analysis of change in this unique habitat. To date, the survey has never been fully repeated, however thanks to the availability of the published information, plans are underway to repeat the survey between 2020 and 2022, under the direction of the Woodland Trust.

Links to my other works

This paper has very close links to the Woodland Survey and related data sets, as described above.

Work undertaken

Before undertaking this work, the information and data from this survey were severely at risk of loss. My work undertaken for this paper involved reconstructing data sets from available information including decoding data and information from archived documents, locating and digitising site locations from archived documents at UKCEH and archived documents belonging to original staff, understanding and clarifying the data collection methods from archived handbooks and interviews with original staff, digitising soil information (from archived laboratory notebooks and other information) and background research into the pinewoods. The paper also contains my analyses of the most frequent 25 plant species and a re-analysis of the diameter distribution of trees. The originator of the survey, R.G.H. Bunce, was the second author and assisted with background information and context of the surveys.

Citations

To date this article has one citation.

3. Wood, Claire M.; Bunce, Robert G.H. (2016) **Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring**. Earth System Science Data, 8 (1). 89-103. https://doi.org/10.5194/essd-8-89-2016

Summary

This survey was originally initiated in response to the threat posed by the discovery of North Sea oil off the coast of Shetland in the 1960s. Concern was expressed that new development would threaten the natural features of Shetland such as the landscape and wildlife. A statistical environmental framework was constructed by the Institute of Terrestrial Ecology on which to select samples for the survey. Vegetation and habitat data were then collected. The Shetland survey was a crucial stage in the development of the standardised methodology of strategic ecological survey. At this time, the statistical sampling methods used were generally not widely used as a method for ecological monitoring. This paper represents the first time the data and results have been publicly reported and made available and provides an opportunity for a repeat survey, which would be an opportunity to enable the changes in the islands, both socio-economic and environmental to be quantified and to be used for developing mitigating policies.

Contribution to field

The data and publication make available the information representing the ground-breaking methods developed for repeatable, long-term monitoring. There are few other data sets on the vegetation and habitats of Shetland, particularly from the pre-oil industry era.

Links to my other works

This paper is significant because it describes the forerunner of the national Countryside Survey and subsequent surveys. The paper also explains how the methods developed.

Work undertaken

Before undertaking this work, the information and data from this survey were severely at risk of loss. My work undertaken for this paper involved reconstructing data sets from available information including interpreting the coded data from archived documents as well as interviews with original staff. The sampling framework was reconstructed from archival material. It was also necessary to locate and digitise site locations from archived documents at UKCEH, as well as those belonging to original staff. It was necessary to clarify the data collection methods from archived handbooks and interviews with original staff, to digitise soil information from archived laboratory notebooks and other information, and to undertake background research into Shetland. The paper also contains my analyses of the most frequent 25 plant species and a species distribution map of the *Calluna/Eriophorum* vegetation group. The originator of the survey, R.G.H. Bunce, was the second author and assisted with background information and context to the surveys.

Citations

To date this article has five citations.

4. Wood, Claire M.; Bunce, Robert G.H.; Norton, Lisa R.; Maskell, Lindsay C.; Smart, Simon M.; Scott, W. Andrew; Henrys, Peter A.; Howard, David C.; Wright, Simon M.; Brown, Michael J.; Scott, Rod J.; Stuart, Rick C.; Watkins, John W. (2018) **Ecological landscape elements: long-term monitoring in Great Britain, the Countryside Survey 1978-2007 and beyond**. Earth System Science Data, 10 (2). 745-763. https://doi.org/10.5194/essd-10-745-2018

Summary

This paper represents the element of the Countryside Survey that involves collecting the habitat and landscape mapping data, collected between 1978 and 2007. The survey design is based on a series of dispersed, stratified, randomly selected 1 km squares from across Britain, which numbered 256 in 1978, 506 in 1990, 569 in 1998 and 591 in 2007. The stratification used was the statistical environmental classification of 1 km squares in Great Britain as described in Bunce et al. (1996b). Detailed information regarding vegetation types and land use was mapped in all five surveys, allowing reporting by defined standard habitat classifications. Additionally, point and linear landscape features, such as trees and hedgerows, are available from all surveys after 1978. From these stratified, randomly located sample squares, the information can be converted into national estimates, with associated error terms.

Contribution to field

The survey as a whole provides a wealth of globally unique ecological data, consisting of an extensive range of measurements at a national scale, covering a time span of 29 years. From an international perspective, Countryside Survey was a pioneer in surveys of its type. The integrated, systematic national monitoring of vegetation species, soils and landscape features across all land uses provided by Countryside Survey was a novel concept, preceding programmes in many other countries particularly in Europe (Dramstad et al., 2002; Hintermann et al., 2002; Ståhl et al., 2011) and beyond (Burton et al., 2014). A need to disseminate information from this large national survey has partly been a driver for the development of a robust data centre in UKCEH.

Links to my other works

The data described in this paper are analysed in the 'uplands' paper (Bunce et al., 2018b). The paper is linked to the Vegetation Paper (Wood et al., 2017) and builds upon the other surveys cited (Papers 1-3).

Work undertaken

My contribution to the paper was managing the major exercise in data rescue for 1978 (Wood et al., 2012), improving data quality for all data sets 1978 - 2007 and preparing documentation for data re-use. I prepared the manuscript and carried out the documentation and publication work. The other authors have all been involved in the delivery of the actual survey at some point in its history. R.G.H. Bunce designed the sampling framework and survey strategy in 1978. R.G.H. Bunce, S. M. Smart, L.C. Maskell, L.R. Norton and D.C. Howard have all been part of the Countryside Survey co-ordination team for at least one survey, with W.A. Stott and P.A. Henrys contributing statistical support. S.M. Wright, M.J. Brown, R.J. Scott, R. Stuart and J.W. Watkins have provided technical and data management support to the survey for several years.

Citations

To date this article has five citations.

5. Wood, Claire M.; Smart, Simon M.; Bunce, Robert G.H.; Norton, Lisa R.; Maskell, Lindsay C.; Howard, David C.; Scott, W. Andrew; Henrys, Peter A. (2017) **Long-term vegetation monitoring in Great Britain - the Countryside Survey 1978-2007 and beyond**. Earth System Science Data, 9 (2). 445-459. https://doi.org/10.5194/essd-9-445-2017

Summary

This paper represents a major element of the Countryside Survey; the vegetation data. It describes the details of the vegetation surveys, sampled from within 1km square survey sites and carried out in 1978, 1990, 1998 and 2007. The plots sample vegetation in land at random (large 200m^2 plots) and also in road verges, hedgerows and boundaries, stream sides, targeted/rarer habitat types, unenclosed habitats and arable margins.

Contribution to field

As stated above, the vegetation plots are part of a set of globally unique ecological data, covering a time span of 29 years.

Links to my other works

The data described in this paper are analysed in the ‘uplands’ paper (Bunce et al., 2018b). The paper is linked to the Landscape Elements paper (Wood et al., 2018b) because the surveys were undertaken contemporaneously, building upon methods from the earlier surveys cited.

Work undertaken

I prepared the manuscript and carried out the documentation (partly building on previous work by S. Smart) and publication work. The other authors have all been involved in the delivery of the actual survey at some point in its history. R.G.H. Bunce designed the sampling framework and survey strategy in 1978. R.G.H. Bunce, S. M. Smart, L.C. Maskell, L.R. Norton and D.C. Howard have all been part of the Countryside Survey co-ordination team for at least one survey, with W.A. Stott and P.A. Henrys contributing statistical support.

Citations

To date this article has eight citations.

6. Wood, Claire M.; Bunce, Robert G.H.; Norton, Lisa R.; Smart, Simon M.; Barr, Colin J. (2018) **Land cover and vegetation data from an ecological survey of `key habitat' landscapes in England, 1992-1993.** Earth System Science Data, 10 (2). 899-918. https://doi.org/10.5194/essd-10-899-2018

Summary

The sampling framework for the national Countryside Survey is not optimised to yield data on rarer or more localised habitats, therefore a survey was commissioned in the 1990s to carry out additional survey work in English landscapes which contained semi-natural habitats that were perceived to be under threat, or which represented areas of concern to the former Department of the Environment (lowland heath, chalk and limestone (calcareous) grasslands, coasts and uplands). The information recorded allowed an assessment of the extent and quality of a range of the habitats defined during the project, which can now be translated into standard UK broad and priority habitat classes (Jackson, 2000; Maddock, 2008). The survey, known as the ‘Key Habitat Survey’, followed a design which was a series of gridded, stratified, randomly selected 1 km squares taken as representative of each of the four landscape types in England, determined from statistical land classification and geological data (‘spatial masks’). A total of 213 of the 1 km square sample sites were surveyed in the summers of 1992 and 1993, with information being collected on vegetation species, land cover, landscape features and land use, applying standardised repeatable methods.

Contribution to field

This paper describes a major data set providing details on the distribution and quality of rarer habitat and landscape types in England. It contributes additional information and value to the long-term monitoring data gathered by the Countryside Survey and provides a valuable baseline against which future ecological changes may be compared, offering the potential for a repeat survey. The survey is not well-known, and has often been overlooked due to the previous inaccessibility of the data and documents. The survey shows how the Countryside Survey methods can be adapted in order to focus on specific habitats or areas of interest.

Links to my other works

The main link is to the Countryside Survey papers and data as ancillary information providing more detail describing rarer/targeted habitats.

Work undertaken

This set of data comprised perhaps the most challenging data rescue exercise of all of the examples presented. The survey was large and complex, and the methods for siting survey plots and sites changed during the survey due to funding restraints. The original survey was completed immediately before staff were busy planning for Countryside Survey in 1998, and so the data were never archived properly or revisited, but largely abandoned. Whilst the paper field sheets were available, much data was stored in defunct Oracle databases at UKCEH in different schemas with no metadata. The first task was to reconstruct the sampling framework used for the survey. Information from old Oracle databases, archived handbooks, paper documents and field sheets, staff and unpublished reports were relied upon heavily to render the data re-usable. I prepared the manuscript and performed the data rescue work, also the boundary type analyses and vegetation species analysis in the paper. The sampling framework and survey strategy was based on methods designed by R.G.H. Bunce, and the field survey was overseen by C.J. Barr.

Citations

This article does not yet have any citations.

7. Bunce, Robert G.H.; Wood, Claire M.; Smart, Simon M. (2018) **The ecology of British upland landscapes. I. Composition of landscapes, habitats, vegetation and species.** Journal of Landscape Ecology, 11 (3). 120-139. https://doi.org/10.2478/jlecol-2018-0015

Summary

This paper is an example of how data collected in the statistically robust, repeatable method outlined in the previous papers may be used to answer questions on national trends in relation to a particular issue, in the case of the British uplands. The paper was stimulated by the lack of adequate statistics regarding the natural and biodiversity characteristics of the British uplands. This paper therefore provides a comprehensive, definitive set of statistics for the British uplands. An overview of the background to the region is first provided, together with some examples of the available figures and a discussion of their limitations. The paper uses a formal structure, with landscapes at the highest level followed by habitats, then vegetation, and finally species, with exact definitions of the categories applied at all levels. The figures are produced from a survey of stratified, random 1 km squares. The tables give comprehensive figures for Great Britain (GB) as a whole, and also England, Wales and Scotland.

Contribution to field

Official figures for upland habitats in Britain are not comprehensive, and moreover are not the product of robust statistical sampling and analysis. This paper fills this gap by presenting an integrated comprehensive picture of the state of the landscape and habitat of the British uplands based on robust data collected in the Countryside Survey in 2007.

Links to my other works

This paper demonstrates how data provided by the Countryside Survey (Wood et al., 2017; 2018) may be used to answer specific questions, in this case an overall series of statistics for the habitats, vegetation and landscapes of the British uplands. The paper links strongly to the Countryside Survey papers.

Work undertaken

My work for this paper involved all of the analyses, comprising the *c*ompilation of figures for upland areas based on the ITE Land Classification (Bunce et al., 2007; Bunce et al., 1996a), the compilation and analysis of figures of Broad Habitats from Countryside Survey 2007 (Brown et al., 2016b), the analysis of grazing data from Countryside Survey, the compilation and analysis of figures for linear features from Countryside Survey 2007 (Brown et al., 2016c), the compilation and analysis of figures using the Countryside Vegetation System Classes and the compilation and analysis of figures regarding upland vegetation species. R.G.H. Bunce and S. Smart provided context and discussion on the results of the analyses.

Journal standing and citations

The Journal of Landscape Ecology is a fully reviewed scientific journal published by the Czech National Chapter of the Association for Landscape Ecology (CZ-IALE). The journal aims to fill a gap in the ecological field scope covered by the European scientific journals, particularly for those produced in Czech Republic but not limited territorially. Themes are focused on landscape-ecological issues. It has an impact factor of 1.1.

To date this article has two citations.

5. Summary and Conclusions

5.1 Contribution of the work overall

My work is ultimately an example of a success story, with a legacy to take forward for future long-term ecological research. There are several conclusions that may be drawn as to the wider context and significance of this UKCEH 'Bunce' data rescue and publication exercise and approach, and the legacy of the surveys and their data. These conclusions are presented in terms of the two concepts of originality of the work introduced in Section 1, firstly as a novel synthesis of existing information, and secondly in adding to knowledge in a way that has not been done before. The lessons learnt from the process are described, together with suggestions for future actions and policies.

5.2 Novel synthesis of existing information

Although the early 'Bunce' data sets theoretically still remained in existence before this rescue work commenced, a high proportion of the information was at serious risk of loss, was not in a usable format and was not accompanied by any kind of coherent descriptive information. The data, as a usable product, did not exist due to a lack of technological tools and skills at the time of data collection, therefore the data related work in relation to the 'Bunce' surveys is novel. Of the eight stages of a research project relating to ecological informatics proposed by Michener and Jones (2012), this work contributes to the majority of stages, namely planning data storage in terms of the (relational) database design, (re-) acquiring and managing data, elements of quality assurance and control in terms of data quality, the production of comprehensive metadata, the long-term archiving of data and metadata in a public repository (providing public access and dissemination) and the publication of all related data and information in peer-reviewed journals, including to some extent, examples of data analyses. In contributing to these stages, the following practical achievements have been completed:

- 27 individual data sets (as listed in Section 3) are now publicly available in a re-usable format, accompanied by detailed metadata. The data and metadata have been published in peer-reviewed journals and are now available for anyone to use, enabling potential repeats of surveys, transparency, accessibility, data use and re-use.
- Using the same procedures, further similar data sets (Adamson et al., 2017; Bunce et al., 2017a; Bunce et al., 2018a; Bunce et al., 2017c) have also been published with detailed metadata (although have not yet been described or peer reviewed in published papers).

In undertaking this synthesis of new information, primary sources such as paper archives, digital archives and personnel from the original work have all been consulted in order to produce comprehensive documentation to describe the newly available data sets. These final outputs (data and metadata) from the surveys are distinct from the work carried out to collect data and carry out the surveys in the first place.

5.3 Adding to knowledge in a way that has not been done before

The work adds to knowledge by providing publicly accessible, quality assured data and metadata that were not previously available, effectively providing 'new' data and information. This raises the prospect of new analyses, of the sort partly demonstrated in the paper regarding the upland habitats of Britain (Paper 7), and to some extent demonstrated in the other six papers. Thanks in part to the availability of the 'new' data sets, the Woodland Survey and the Pinewood Survey are being repeated, starting in 2020. The repeat will yield a valuable new resource, allowing the exploration of many ecological issues in light of new drivers and issues such as climate change and biodiversity loss. The opportunity to be able to study changes over time is a major incentive for repeating a survey and thus provides clear justification for carrying out this type of rescue work as part of a long-term monitoring programme. Such analysis of long-term change is especially important in current dynamic times.

The approach to data rescue and publication was also novel and was carried out alongside major developments in the provision of long-term data storage and dissemination activities provided by the Natural Environment Research Council (NERC) Environment Information Data Centre. The earlier papers and data deposits served as case studies, helping to inform and direct best practice within the data centre in terms of setting out requirements for metadata, supporting documentation and deposit procedures and in turn, improved as the guidance provided by the data centre was improved.

5.4 Lessons learnt from the rescue exercise

The rescue exercise overall provides a framework for others to follow, where potential for rescuing similar data may arise. It highlights the type of original survey documents, including data recording sheets, field handbooks, code sheets, notes, maps and photographs, required to document a complex ecological data set, describes some of the technical challenges to be overcome, and provides examples of fully documented data sets. Many other data sets are likely to exist and could follow similar rescue procedures as those described here.

By following this process for legacy data, it also highlights the type of pitfalls which may occur when data are not factored into the workflow at the start of an ecological field research project, in terms of collection, database design and management, through to analysis. As new technologies for the collection, handling and analysis of data continue to emerge, it is more important than ever to plan the data flow throughout a project carefully. A specific pitfall might include the loss or potential loss of key information such as descriptions for coded data. Even once securely held in a database, it is not always wise to follow absolute best practice in relational database design in this case; it is often advantageous to store code descriptions alongside key data in the same table. Where there are several tables of this type in a database, the descriptive tables are easily confused or lost. Ecological data are unique and sometimes needs dealing with their own set of rules and procedures. Another pitfall is not giving thought to how data will be stored and retrieved from a database at the point of collection. Data entered into a spreadsheet in an unstructured way can potentially be as useless as no data at all. The development of electronic data capture methods in the field facilitates data workflow from the field to final storage, however, again, thought needs to be given to this when planning projects.

At an institutional level, a range of lessons have also been learnt. UKCEH have demonstrated that it is advantageous for institutes to maintain contact with former staff through a fellowship scheme. Extending the amount of time former staff are available to question also extends the window of time in which a data set may be rescued. In the case of complex surveys such as these, data rescue must be achieved before it is too late and knowledge is lost forever. The achievements outlined in this document highlight the value of libraries and archives, curating the essential resources required to describe legacy data sets. Documents need curating and archiving properly for long-term storage. This work proves how valuable this type of information can be.

Overall, the main lesson to be learnt is identifying clearly the importance of planning current and future projects, in order to avoid either the loss of data (collected at considerable expense), or alternatively, having to undertake complex rescue exercises. The importance of good data management practices is often ignored, particularly by scientists. Improvements in policies and procedures have been achieved, particularly in the past decade (Tenopir et al., 2020) but there is still a long way to go, and the need for staff skilled in data management in still overlooked. This work demonstrates that one of the most costly aspects of managing data is the need for data restructuring, and there is still little recognition of this within project planning.

5.5 The future

5.5.1 The 'Bunce' data

Looking to the future, the legacy of the 'Bunce' UKCEH data sets is their ability to be repeated, thus providing a solid baseline and integrated platform for monitoring land use in the UK countryside as a whole, and also specific habitats (woodlands, 'Key Habitats') and regions (Cumbria, Shetland). The data sets provide a method of delivering scientifically robust information regarding vegetation, soils and freshwaters to identify and tackle the many issues (both known and currently unknown) facing the environment, such as biodiversity loss, habitat degradation and declining soil carbon. Surveys that have not yet been repeated, such as the Shetland, Pinewoods and 'Key Habitat' surveys, are in a position to follow the lead of the Countryside Survey, which has clearly demonstrated how repeat surveys can develop into a valuable long-term monitoring framework. The Woodland Survey, is also about to be repeated, for the third time, and will yield important policy related results.

In terms of potentially tackling a future data rescue project, the data sets provide an example of how a similar result may be achieved with other data sets, and also provide examples of data and metadata packages which potentially could be used as a template for other data publications.

5.5.2 Publicly funded data

In more general terms, it seems that newly collected data assets, particularly in publicly-funded institutes like UKCEH, have a higher chance of survival than in the past now that new procedures are in place for planning and storing data. Within NERC, the data policy requires researchers in receipt of NERC grants to offer their data to a NERC or other relevant long-term repository. This is being increasingly enforced through sanctions to grant-holders for non-compliance. Besides EIDC, there are also many other national and international long-term repositories available. Although some are more robust and better quality than others, there is no excuse not to deposit data for the long-term. 'Forward-thinking ecologists will organise and archive data for posterity, publicly share their data, and participate in collaborations that address large-scale questions' (Hampton et al., 2013).

5.5.3 EIDC

Regarding EIDC, the data centre infrastructure is now established, and it is becoming a mainstream requirement for NERC-funded researchers to deposit their data there. Future practical challenges may lie in coping with increased numbers of deposits and downloads and curating increasing numbers of data sets for the long-term. The Head of EIDC recognises the need for the data centre to go beyond being a repository for data sets (particularly for NERC grant holders). In particular, EIDC should become more integrated into the day to day science of UKCEH and should facilitate large scale analytics – moving away from the idea of researchers downloading local copies of data sets to work with, and embracing the concept of 'bringing the code to the data' using online analytical tools such as 'DataLabs' (https://datalab.datalabs.ceh.ac.uk/). This kind of integrative approach can help gain the maximum benefit from the disparate collections of data gathered by UKCEH (J.W. Watkins, personal communication, 29/9/2020).

5.5.4 Recognition for data management work

The future for staff working with data and data management tasks seems more positive in terms of recognition and reward. Traditionally, academia has not adequately rewarded the time and energy required to comprehensively document a database or information product, preferring instead to assign a much higher value to funded grants and peer reviewed publications (Michener, 2006). Now, the ability to publish data sets gaining citations with digital object identifiers, and publish data papers in peer-reviewed journals provides an incentive for (data) scientists to archive and document their data appropriately. This will result in significant gains for both the current research community and scientists for decades to come (Callaghan et al., 2012). 'To fully take advantage of scientific opportunities available in the information age, ecologists must treat data as an enduring product of research and not just as a precursor to publications' (Hampton et al., 2013).

5.5.5 Legacy data

Unfortunately, the prospects for legacy data sets at risk do not seem assured, even within UKCEH. Resourcing data rescue efforts can be expensive and time consuming, especially in relation to personnel time. Also, few want to 'poke around musty archives for heritage data captured using yesterday's technology' (Griffin, 2017).

UKCEH does have several factors in its favour in this regard. Firstly, it has a long institutional memory. Despite site moves and staff retirements, many documents have fortunately been kept intact in the archives. UKCEH has a 'fellowship' scheme, which means retired staff remain available for longer, to provide information and memories regarding previous work. Also, the Informatics Liaison Network, set up in 2006, exists to improve data management and data management practices within UKCEH, and has provided some time to facilitate the 'Bunce' data rescue to take place. Undertaking repeats of the 'Bunce' surveys increasingly demonstrates the value for money of the cost and effort invested in rescuing the irreplaceable data sets.

However, despite the huge improvements in the ability to manage and store data outputs from contemporary research, particularly in the last 10 or 15 years, there are still many old boxes of nationally important information sitting on shelves, or on obsolete storage media, with no particular champion to move their rescue forward (Hampton et al., 2013). Within UKCEH, on the whole, there is a general lack of interest in legacy data sets, and little support for rescuing them and sadly, time is running out.

5.5.6 Data management planning

The amount of work and resource that goes into a data rescue effort, although worthwhile, can be avoided in future if work is planned properly from the outset, with a view to the long-term storage and reusability of data outputs. Planning should emphasise the importance of documenting information and workflows throughout a project, in order to ensure transferability. Newly collected data assets, particularly in publicly-funded institutes like UKCEH, have a higher chance of survival than in the past now that new procedures are in place for planning and storing data; it is recognised that the time and effort put into collecting data should not be wasted. This is particularly the case in relation to long-term monitoring, as it expensive and complex to achieve, therefore the data needs managing well.

NERC and UKCEH recognise this, and have put effort and resource into ensuring that data are considered at the outset of a project, through use of Data Management Plans (for example UKRI-NERC (2020)) and through the establishment of an Informatics Liaison Network, providing advice on best practice in data management throughout UKCEH.

6. References

Adamson, J. K., Benefield, C. B., Elfyn Hughes, W., & Wood, C. M. (2017). Terrestrial habitat and vegetation data from the marginal uplands of Cumbria, 1978. https://doi.org/10.5285/a4f2d52b-1515-434e-bfc3-83b3a53be1c5: NERC Environmental Information Data Centre.

AGI. (2020a). AGI GEMINI Update Retrieved 30/11/2020, from https://www.agi.org.uk/40-gemini/

AGI. (2020b). UK GEMINI standard and INSPIRE implementing rules Retrieved 30/11/2020, from https://www.agi.org.uk/gemini/1037

Alves, C., Castro, J. A., Ribeiro, C., Honrado, J. P., & Lomba, A. (2018). *Research data management in the field of Ecology: an overview.* Paper presented at the International Conference on Dublin Core and Metadata Applications.

Arrouays, D., Leenaars, J. G., Richer-de-Forges, A. C., Adhikari, K., Ballabio, C., Greve, M., Grundy, M., Guerrero, E., Hempel, J., & Hengl, T. (2017). Soil legacy data rescue via GlobalSoilMap and other international and national initiatives. *GeoResJ, 14*, 1-19. doi: https://doi.org/10.1016/j.grj.2017.06.001

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly, 32*(4), 399-418. doi: https://doi.org/10.1016/j.giq.2015.07.006

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Fuller, R. M., Furse, M. T., Gillespie, M. K., Groom, G. B., Hallam, C. J., Hornung, M., Howard, D. C., & Ness, M. J. (1993). Countryside Survey 1990: main report. (Countryside 1990 vol.2): London: Department of the Environment.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2016a). Landscape area data 1990 [Countryside Survey]. https://doi.org/10.5285/94f664e5-10f2-4655-bfe6-44d745f5dca7: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2016b). Landscape linear feature data 1990 [Countryside Survey]. https://doi.org/10.5285/311daad4-bc8c-485a-bc8a-e0d054889219: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2016c). Landscape point feature data 1990 [Countryside Survey]. https://doi.org/10.5285/1481bc63-80d7-4d18-bcba-8804aa0a9e1b: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., & Wright, S. M. (2016d). Landscape area data 1998 [Countryside Survey]. https://doi.org/10.5285/1e050028-5c55-42f4-a0ea-c895d827b824: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., & Wright, S. M. (2016e). Landscape linear feature data 1998 [Countryside Survey].

https://doi.org/10.5285/8aaf6f8c-c245-46bb-8a2a-f0db012b2643: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., & Wright, S. M. (2016f). Landscape point feature data 1998 [Countryside Survey]. https://doi.org/10.5285/ed10944f-40c8-4913-b3f5-13c8e844e153: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Cummins, R. P., Hallam, C. J., Hornung, M., & Wood, C. M. (2017). Habitat and vegetation data from an ecological survey of terrestrial key habitats in England, 1992-1993. doi:10.5285/7aefe6aa-0760-4b6d-9473-fad8b960abd4: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Smart, S. M., Stuart, R. C., & Wood, C. M. (2014a). Countryside Survey 1990 vegetation plot data. https://doi.org/10.5285/26e79792-5ffc-4116-9ac7-72193dd7f191 NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Shield, E. R., Smart, S. M., Stuart, R. C., Watkins, J. W., & Wood, C. M. (2014b). Countryside Survey 1998 vegetation plot data. https://doi.org/10.5285/07896bb2-7078-468c-b56d-fb8b41d47065 NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., & Wright, S. M. (2016g). Landscape area data 1984 [Countryside Survey]. https://doi.org/10.5285/b656bb43-448d-4b2c-aade-7993aa243ea3: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., & Wright, S. M. (2016h). Landscape linear feature data 1984 [Countryside Survey]. https://doi.org/10.5285/a3f5665c-94b2-4c46-909e-a98be97857e5: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., & Wright, S. M. (2016i). Landscape point feature data 1984 [Countryside Survey]. https://doi.org/10.5285/124b872e-036e-4dd3-8316-476b5f42c16e: NERC Environmental Information Data Centre.

Barr, C. J., Bunce, R. G. H., Smart, S. M., & Whittaker, H. A. (2014c). Countryside Survey 1978 vegetation plot data. https://doi.org/10.5285/67bbfabb-d981-4ced-b7e7-225205de9c96 NERC Environmental Information Data Centre.

Barr, C. J., & Gillespie, M. K. (2000). Estimating hedgerow length and pattern characteristics in Great Britain using Countryside Survey data. *Journal of Environmental Management, 60*(1), 23-32. doi: https://doi.org/10.1006/jema.2000.0359

Barr, C. J., Howard, D. C., Bunce, R. G. H., Gillespie, M. K., & Hallam, C. J. (1991). Changes in hedgerows in Britain between 1984 and 1990. Contract report to the Department Of The Environment. Grange-over-Sands: Institute of Terrestrial Ecology.

Barr, C. J., & Parr, T. W. (1994). Hedgerows: Linking Ecological Research and Countryside Policy. In T. A. Watt & G. P. Buckley (Eds.), *Hedgerow Management and Nature Conservation* University of London: Wye College Press.

BBC. (2020). Help needed to rescue UK's old rainfall records Retrieved 22/09/2020, 2020, from https://www.bbc.co.uk/news/science-environment-52040822

Berkley, C., Bowers, S., Jones, M. B., Madin, J. S., & Schildhauer, M. (2009). Improving data discovery for metadata repositories through semantic search *2009 International*

Conference on Complex, Intelligent and Software Intensive Systems (pp. 1152-1159). https://doi.org/10.1109/CISIS.2009.122: IEEE.
Bowser, C. J. (1986). Historic data sets: lessons from the past, lessons for the future. In W. Michener (Ed.), *Research data management in the ecological sciences* (pp. 155-179). South Carolina: University of South Carolina Press.
Braun-Blanquet, J. (1932). *Plant sociology. The study of plant communities*. New York: McGraw-Hill.
Bretherton, F. P., & Singley, P. T. (1994). Metadata: A user's view *Seventh International Working Conference on Scientific and Statistical Database Management* (pp. 166-174). https://doi.org/10.1109/SSDM.1994.336950: IEEE.
British Oceanographic Data Centre. (2020). British Oceanographic Data Centre Retrieved 27/11/2020, from https://www.bodc.ac.uk/
Brown, M. J., Bunce, R. G. H., Carey, P. D., Chamberlain, P. M., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2016a). Landscape point feature data 2007 [Countryside Survey]. https://doi.org/10.5285/55dc5fd7-d3f7-4440-b8a7-7187f8b0550b: NERC Environmental Information Data Centre.
Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2016b). Landscape area data 2007 [Countryside Survey]. https://doi.org/10.5285/bf189c57-61eb-4339-a7b3-d2e81fdde28d: NERC Environmental Information Data Centre.
Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2016c). Landscape linear feature data 2007 [Countryside Survey]. https://doi.org/10.5285/e1d31245-4c0a-4dee-b36c-b23f1a697f88 NERC Environmental Information Data Centre.
Bunce, R. (1968). Biomass and production of trees in a mixed deciduous woodland: I. Girth and height as parameters for the estimation of tree dry weight. *The Journal of Ecology*, 759-775. doi: https://doi.org/10.2307/2258105
Bunce, R. G. H. (1981). British woodlands in an European context *In: Forest and woodland ecology: an account of research being done in ITE, edited by F.T.Last & A.S.Gardiner, 7-11. (ITE symposium no.8). Cambridge: Institute of Terrestrial Ecology.*
Bunce, R. G. H. (1982). *A field key for classifying British woodland vegetation. Part 1*. Cambridge Institute of Terrestrial Ecology.
Bunce, R. G. H. (1989). *A field key for classifying British woodland vegetation. Part 2*. London HMSO.
Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D., & Lane, A. M. J. (1990). ITE Land Classification of Great Britain 1990. https://doi.org/10.5285/ab320e08-faf5-48e1-9ec9-77a213d2907f: NERC Environmental Information Data Centre.
Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D., & Lane, A. M. J. (1998). ITE Land Classification of Great Britain 1998. https://doi.org/10.5285/971671a6-98b4-4d80-b165-21dace7373b9: NERC Environmental Information Data Centre.
Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D., & Scott, A. (2007). ITE Land Classification of Great Britain 2007. https://doi.org/10.5285/5f0605e4-aa2a-48ab-b47c-bf5510823e8f: NERC Environmental Information Data Centre.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., & Lane, A. M. J. (1996a). ITE Merlewood Land Classification of Great Britain. *Journal of Biogeography, 23*, 625-634. doi: https://doi.org/10.1111/j.1365-2699.1996.tb00023.x

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., & Howard, D. C. (1996b). The ITE Land Classification: providing an environmental stratification of Great Britain. *Environmental Monitoring and Assessment, 39*, 39-46. doi: https://doi.org/10.1007/978-94-009-1653-1_5

Bunce, R. G. H., Barr, C. J., Shaw, M. W., Bunce, F., Bassett, P., Quarmby, S., Robertson, S. M. C., Bowen, W. J. A., Morris, D., Millar, A., Brown, A. H. F., Horrill, A. D., Dodd, B. A., & Wood, C. M. (2017a). Ecological survey of the Spey Valley birch woods, 1973. https://doi.org/10.5285/c84961e3-b9dc-4c92-b316-36295b8a3330: NERC Environmental Information Data Centre.

Bunce, R. G. H., Barr, C. J., & Whittaker, H. A. (1981). An integrated system of land classification *Annual Report of the Institute of Terrestrial Ecology* (Vol. 1980, pp. 28-33.). Huntingdon: Institute of Terrestrial Ecology.

Bunce, R. G. H., & Bassett, P. A. (2015). Land Classification of Shetland 1974. https://doi.org/10.5285/f1b3179e-b446-473d-a5fb-4166668da146 NERC Environmental Information Data Centre.

Bunce, R. G. H., Bassett, P. A., & Wood, C. M. (2015a). Terrestrial habitat, vegetation and soil data from Shetland, 1974. https://doi.org/10.5285/06fc0b8c-cc4a-4ea8-b4be-f8bd7ee25342: NERC Environmental Information Data Centre.

Bunce, R. G. H., Brown, A. H. F., Shaw, M. W., White, E. J., Gardiner, A., Procter, J. E. A., Birkett, J. S., McAllister, J., Smith, C., Parrington, J., Wrigley, J., & Wood, C. M. (2018a). Tree, vegetation and soil information from an ecological survey of semi-natural woodlands in the English Lake District, 1969: NERC Environmental Information Data Centre.

Bunce, R. G. H., Carey, P. D., Maskell, L. C., Norton, L. R., Scott, R. J., Smart, S. M., & Wood, C. M. (2014). Countryside Survey 2007 vegetation plot data. https://doi.org/10.5285/57f97915-8ff1-473b-8c77-2564cbd747bc NERC Environmental Information Data Centre.

Bunce, R. G. H., & Jeffers, J. N. R. (1977). *Native pinewoods of Scotland*. Cambridge: Institute of Terrestrial Ecology.

Bunce, R. G. H., Parr, T. W., Ullyett, J., Hornung, M., Gerard, F., Bull, R., Cox, R., & Brown, N. J. (2017b). Spatial masks for calcareous, coastal, upland and lowland heath landscapes in England [Key Habitats 1992-93]. https://doi.org/10.5285/dc583be3-3649-4df6-b67e-b0f40b4ec895: NERC Environmental Information Data Centre.

Bunce, R. G. H., & Shaw, M. W. (1972). Classifying woodlands for nature conservation. *Forestry and British Timber, 1*, 23-25.

Bunce, R. G. H., & Shaw, M. W. (1973). A standardised method for ecological survey. *Journal of Environmental Management, 1*(239-258), 239-258.

Bunce, R. G. H., Shaw, M. W., & Wood, C. M. (2015b). Habitat, vegetation, tree and soil data from Native Pinewoods in Scotland, 1971 *NERC Environmental Information Data Centre*. https://doi.org/10.5285/56a48373-771c-4d4a-8b5a-45ef496c6e55: NERC Environmental Information Data Centre.

Bunce, R. G. H., & Smith, R. S. (1978). An ecological survey of Cumbria *Structure Plan Working Paper, 4* (Vol. Structure Plan Working Paper, 4). Kendal Cumbria County Council and Lake District Special Planning Board.

Bunce, R. G. H., Smith, R. S., Booy, M., & Wood, C. M. (2017c). Terrestrial habitat, vegetation and soil data from Cumbria, England, 1975. https://doi.org/10.5285/e96b909d-d52e-4b36-bc51-905ece420794: NERC Environmental Information Data Centre.
Bunce, R. G. H., Wood, C., Henrys, P., Smart, S., Howard, D., & Barr, C. J. (2016). Landscape area data 1978 [Countryside Survey].
Bunce, R. G. H., Wood, C. M., & Smart, S. M. (2018b). The Ecology of British Upland landscapes. I. Composition of landscapes, habitats, vegetation and species. *Journal of Landscape Ecology, 11*(3). doi: https://doi.org/10.2478/jlecol-2018-0015
Burton, A. C., Huggard, D., Bayne, E., Schieck, J., Sólymos, P., Muhly, T., Farr, D., & Boutin, S. (2014). A framework for adaptive monitoring of the cumulative effects of human footprint on biodiversity. *Environmental monitoring and assessment, 186*(6), 3605-3617. doi: https://doi.org/10.1007/s10661-014-3643-7
Bystrak, D. (1981). The North American breeding bird survey (Vol. 6). Lawrence, Kansas: Cooper Ornithological Society, Allen Press.
Caldwell, P. (2012). Tide gauge data rescue. *Proceedings of The Memory of the World in the Digital age: Digitization and Preservation*, 134-149.
Callaghan, S., Donegan, S., Pepler, S., Thorley, M., Cunningham, N., Kirsch, P., Ault, L., Bell, P., Bowie, R., & Leadbetter, A. (2012). Making data a first class scientific output: Data citation and publication by NERC's environmental data centres. *INTERNATIONAL JOURNAL OF DIGITAL CURATION, 7*(1). doi: DOI: https://doi.org/10.2218/ijdc.v7i1.218
Carey, P. D., Wallis, S., Chamberlain, P. M., Cooper, A., Emmett, B. A., Maskell, L. C., McCann, T., Murphy, J., Norton, L. R., Reynolds, B., Scott, W. A., Simpson, I. C., Smart, S. M., & Ullyett, J. M. (2008). Countryside Survey: UK Results from 2007. Lancaster: NERC/Centre for Ecology & Hydrology.
Carpenter, S. R., Chisholm, S. W., Krebs, C. J., Schindler, D. W., & Wright, R. F. (1995). Ecosystem experiments. *Science, 269*(5222), 324-327. doi: https://doi.org/10.1126/science.269.5222.324
Caughlan, L., & Oakley, K. L. (2001). Cost considerations for long-term ecological monitoring. *Ecological indicators, 1*(2), 123-134. doi: https://doi.org/10.1016/S1470-160X(01)00015-2
CEDA. (2020). The CEDA Archive, 27/11/2020, from http://data.ceda.ac.uk/badc
Cherrill, A., & McClean, C. (1999). The reliability of 'Phase 1' habitat mapping in the UK: the extent and types of observer bias. *Landscape and urban planning, 45*(2-3), 131-143. doi: https://doi.org/10.1016/S0169-2046(99)00027-4
Clapham, A. R., Tutin, T. G., & Warburg, E. F. (1952). *Flora of the British Isles.* Cambridge: Cambridge University Press.
Codd, E. (1970). A relational model of data for large relational databases. *Communications of the ACM, 13*(2), 377-387.
Collins, E. (2011). Use and impact of UK research data centres. *INTERNATIONAL JOURNAL OF DIGITAL CURATION, 6*(1). doi: https://doi.org/10.2218/ijdc.v6i1.169
Corney, P. M., Le Duc, M. G., Smart, S. M., Kirby, K. J., Bunce, R. G. H., & Marrs, R. H. (2004). The effect of landscape-scale environmental drivers on the vegetation composition of British woodlands. *Biological Conservation, 120*(4), 491-505. doi: https://doi.org/10.1016/j.biocon.2004.03.022

Costello, M. J., Michener, W. K., Gahegan, M., Zhang, Z.-Q., & Bourne, P. E. (2013). Biodiversity data should be published, cited, and peer reviewed. *Trends in Ecology & Evolution, 28*(8), 454-461. doi: https://doi.org/10.1016/j.tree.2013.05.002

Critchley, C., Wilson, L., Mole, A., Norton, L., & Smart, S. (2013). A functional classification of herbaceous hedgerow vegetation for setting restoration objectives. *Biodiversity and conservation, 22*(3), 701-717. doi: https://doi.org/10.1007/s10531-013-0440-5

DCC. (2020). Digital Curation Manual Retrieved 30/11/2020, from https://www.dcc.ac.uk/sites/default/files/documents/resource/curation-manual/chapters/file-formats/file-formats.pdf

Diwakar, A. S., Kulkarni, A. A., & Talwai, P. (2008). Rescue and preservation of climate data by extraction and digitization from autographic weather charts using image processing tools *Proc. 8th WSEAS International Conference on Signal, Speech and Image Processing (SSIP'08), World Scientific and Engineering Academy and Society (WSEAS)* (pp. 23-25). https://www.researchgate.net/profile/Anil_Kulkarni/publication/262287157_Rescue_and_preservation_of_climate_data_by_extraction_and_digitization_from_autographic_weather_charts_using_image_processing_tools/links/5ddf8b6b4585159aa44e9114/Rescue-and-preservation-of-climate-data-by-extraction-and-digitization-from-autographic-weather-charts-using-image-processing-tools.pdf.

Downs, R. R., & Chen, R. S. (2017). Curation of scientific data at risk of loss: Data rescue and dissemination. *Curating research data. Volume one, Practical strategies for your digital repository* Retrieved 17/11/2020, from https://academiccommons.columbia.edu/doi/10.7916/D8W09BMQ

Dramstad, W. E., Fjellstad, W., Strand, G.-H., Mathiesen, H. F., Engan, G., & Stokland, J. N. (2002). Development and implementation of the Norwegian monitoring programme for agricultural landscapes. *Journal of Environmental management, 64*(1), 49-63. doi: https://doi.org/10.1006/jema.2001.0503

EIDC. (2020a). Natural England natural capital maps Retrieved 30/11/2020, from https://eip.ceh.ac.uk/naturalengland-ncmaps

EIDC. (2020b). Preparing your datasets for deposit Retrieved 30/11/2020, from https://eidc.ac.uk/deposit/preparingData.

EIDC. (2020c). Suitable formats for data and supporting documentation Retrieved 30/11/2020, from https://eidc.ac.uk/deposit/suitableFormats

EIDC. (2020d). Supporting documentation Retrieved 30/11/2020, from https://eidc.ac.uk/deposit/supportingDocumentation

Elliott, J. (1990). The need for long-term investigations in ecology and the contribution of the Freshwater Biological Association. *Freshwater Biology, 23*(1), 1-5. doi: https://doi.org/10.1111/j.1365-2427.1990.tb00250.x

Ellison, A. M. (2010). Repeatability and transparency in ecological research. *Ecology, 91*(9), 2536-2539.

Emmett, B. E., & GMEP team. (2017). Glastir Monitoring & Evaluation Programme. Final Report to Welsh Government - Executive Summary (Contract reference: C147/2010/11). Bangor: NERC/Centre for Ecology & Hydrology.

ESRI. (2020). Field Operations Retrieved 27/11/2020, from https://www.esri.com/en-us/arcgis/products/field-operations/overview

European Commission. (2020). INSPIRE KNOWLEDGE BASE Retrieved 30/11/2020, from https://inspire.ec.europa.eu/auto-created-vocabulary-2/guidance-document

Garbutt, R., & Sparks, T. (2002). Changes in the botanical diversity of a species rich ancient hedgerow between two surveys (1971–1998). *Biological conservation, 106*(2), 273-278. doi: https://doi.org/10.1016/S0006-3207(01)00253-1
Gibney, E., & Van Noorden, R. (2013). Scientists losing data at a rapid rate. *Nature News*. doi: https://doi.org/10.1038/nature.2013.14416
Gill, P., & Dolan, G. (2015). Originality and the PhD: what is it and how can it be demonstrated? *Nurse researcher, 22*(6). doi: https://doi.org/10.7748/nr.22.6.11.e1335
Griffin, E. (2017). Rescue old data before it's too late. *Nature News, 545*(7654), 267. doi: https://doi.org/10.1038/545267a
Guptill, S. C., & Morrison, J. L. (2013). *Elements of spatial data quality*. Oxford: Elsevier.
Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., Duke, C. S., & Porter, J. H. (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment, 11*(3), 156-162. doi: https://doi.org/10.1890/120103
Hawkins, S., Firth, L., McHugh, M., Poloczanska, E., Herbert, R., Burrows, M., Kendall, M., Moore, P. J., Thompson, R., & Jenkins, S. (2013). Data rescue and re-use: recycling old information to address new policy concerns. *Marine Policy, 42*, 91-98. doi: https://doi.org/10.1016/j.marpol.2013.02.001
Heal, O. (1968). The International Biological Programme Project at Moor House *Merlewood Research and Development Paper* (Vol. 68/2). Grange over Sands: Nature Conservancy.
Hearn, S. M., Healey, J. R., McDonald, M. A., Turner, A. J., Wong, J. L. G., & Stewart, G. B. (2011). The repeatability of vegetation classification and mapping. *Journal of Environmental Management, 92*(4), 1174-1184. doi: https://doi.org/10.1016/j.jenvman.2010.11.021
Henrys, P., Bee, E., Watkins, J., Smith, N., & Griffiths, R. (2015). Mapping natural capital: optimising the use of national scale datasets. *Ecography, 38*(6), 632-638. doi: https://doi.org/10.1111/ecog.00402
Henrys, P., Rowe, E. C., Evans, C. D., Emmett, B. A., Smart, S. M., & Butler, A. (2013). MultiMOVE: A Package of niche models for British vegetation: NERC Environmental Information Data Centre.
Hill, A., Seyfrit, C., & Danner, M. (1998). Oil development and social change in the Shetland Islands 1971-1991. . *Impact Assessment and Project Appraisal 16*, 15-25. doi: https://doi.org/10.1080/14615517.1998.10590183
Hill, L., Hector, A., Hemery, G., Smart, S., Tanadini, M., & Brown, N. (2017). Abundance distributions for tree species in Great Britain: a two-stage approach to modeling abundance using species distribution modeling and random forest. *Ecology and Evolution, 7*(7), 1043-1056. doi: https://doi.org/10.1002/ece3.2661
Hill, M. O., Bunce, R. G. H., & Shaw, M. W. (1975). Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland. *Journal of Ecology, 63*, 597-613. doi: https://doi.org/10.2307/2258738
Hill, M. O., Mountford, J. O., Roy, D. B., & Bunce, R. G. H. (1999). *Ellenberg's indicator values for British plants*. Huntingdon Institute of Terrestrial Ecology.
Hintermann, U., Weber, D., Zangger, A., & Schmill, J. (2002). Biodiversity Monitoring in Switzerland, BDM – Interim Report *Series No. 342.* Berne.

Jackson, D. (2000). Guidance on the Interpretation of the Biodiversity Broad Habitat Classification (Terrestrial and Freshwater Types): Definitions and the Relationship with Other Habitat Classifications (JNCC Report, No 307) (Vol. Report no. 307, pp. 73pp). Published online, http://jncc.defra.gov.uk/page-2433: JNCC.

Jisc. (2020). JISC Retrieved 30/11/2020, from https://www.jisc.ac.uk/

Jones, M. B., Berkley, C., Bojilova, J., & Schildhauer, M. (2001). Managing scientific metadata. *IEEE Internet Computing, 5*(5), 59-68. doi: https://doi.org/10.1109/4236.957896

Kervin, K., Michener, W., & Cook, R. (2013). Common errors in ecological data sharing. *Journal of eScience Librarianship, 2*(2), 3-16. doi: https://doi.org/10.7191/jeslib.2013.1024

Kettlewell, P., Easey, J., Stephenson, D., & Poulton, P. (2006). Soil moisture mediates association between the winter North Atlantic Oscillation and summer growth in the Park Grass Experiment. *Proceedings of the Royal Society B: Biological Sciences, 273*(1590), 1149-1154. doi: https://doi.org/10.1098/rspb.2005.3428

Kimberley, A., Blackburn, G. A., Whyatt, J. D., Kirby, K., & Smart, S. M. (2013). Identifying the trait syndromes of conservation indicator species: how distinct are British ancient woodland indicator plants from other woodland species? *Applied Vegetation Science, 16*(4), 667-675. doi: https://doi.org/10.1111/avsc.12047

Kimberley, A., Blackburn, G. A., Whyatt, J. D., & Smart, S. M. (2016). How well is current plant trait composition predicted by modern and historical forest spatial configuration? *Ecography, 39*(1), 67-76. doi: https://doi.org/10.1111/ecog.01607

Kirby, K., Smart, S., Black, H., Bunce, R., Corney, P., & Smithers, R. (2005). Long-term ecological change in British woodland (1971–2001) *English Nature Research Reports*. Peterborough.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., & Shaw, M. W. (2013a). Woodlands survey flora data 1971-2001. https://doi.org/10.5285/2d023ce9-6dbe-4b4f-a0cd-34768e1455ae: NERC Environmental Information Data Centre.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., & Shaw, M. W. (2013b). Woodlands survey site information 1971-2001. https://doi.org/10.5285/d6409d40-58fe-4fa7-b7c8-71a105b965b4: NERC Environmental Information Data Centre.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., & Shaw, M. W. (2013c). Woodlands survey soil data 1971-2001. https://doi.org/10.5285/fb1e474d-456b-42a9-9a10-a02c35af10d2: NERC Environmental Information Data Centre.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., & Shaw, M. W. (2013d). Woodlands survey tree diameter data 1971-2001. https://doi.org/10.5285/4d93f9ac-68e3-49cf-8a41-4d02a7ead81a: NERC Environmental Information Data Centre.

Knockaert, C., Tyberghein, L., Goffin, A., Vanhaecke, D., Ong'anda, H., Wakwabi, E. O., & Mees, J. (2019). Biodiversity data rescue in the framework of a long-term Kenya-Belgium cooperation in marine sciences. *Scientific data, 6*(1), 1-6. doi: https://doi.org/10.1038/s41597-019-0092-8

Lawrence, B., Jones, C., Matthews, B., Pepler, S., & Callaghan, S. (2011). Citation and peer review of data: Moving towards formal data publication. *International Journal of Digital Curation, 6*(2), 4-37. doi: https://doi.org/10.2218/ijdc.v6i2.205

Likens, G., & Lindenmayer, D. (2018). *Effective ecological monitoring*. Clayton, Australia: CSIRO publishing.

Likens, G. E. (1985). *An ecosystem approach to aquatic ecology. Mirror Lake and its environment*. New York: Springer-Verlag.

Likens, G. E. (1992). The ecosystem approach: its use and abuse. *Excellence in ecology, 3*, VII-XXIV.

Lindenmayer, D. B., & Likens, G. E. (2010). The science and application of ecological monitoring. *Biological conservation, 143*(6), 1317-1328. doi: https://doi.org/10.1016/j.biocon.2010.02.013

Lindenmayer, D. B., Likens, G. E., Andersen, A., Bowman, D., Bull, C. M., Burns, E., Dickman, C. R., Hoffmann, A. A., Keith, D. A., & Liddell, M. J. (2012). Value of long-term ecological studies. *Austral Ecology, 37*(7), 745-757. doi: https://doi.org/10.1111/j.1442-9993.2011.02351.

Maddock, A. (2008). UK Biodiversity Action Plan; Priority Habitat Descriptions *UK Biodiversity Action Plan, 94pp*. Published online, http://jncc.defra.gov.uk/PDF/UKBAP_PriorityHabitatDesc-Rev2011.pdf.

Magurran, A. E., Baillie, S. R., Buckland, S. T., Dick, J. M., Elston, D. A., Scott, E. M., Smith, R. I., Somerfield, P. J., & Watt, A. D. (2010). Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends in ecology & evolution, 25*(10), 574-582. doi: https://doi.org/10.1016/j.tree.2010.06.016

Marrs, R., Kirby, K., Le Duc, M., McAllister, H., Smart, S., Oksanen, J., Bunce, R., & Corney, P. (2013). Native dominants in British woodland-a potential cause of reduced species-richness? *New Journal of Botany, 3*(3), 156-168. doi: https://doi.org/10.1179/2042349713Y.0000000028

Maskell, L. C., Smart, S. M., Bullock, J. M., Thompson, K., & Stevens, C. J. (2010). Nitrogen deposition causes widespread loss of species richness in British habitats. *Global Change Biology, 16*(2), 671-679. doi: https://doi.org/10.1111/j.1365-2486.2009.02022.x

McCollin, D., Jackson, J., Bunce, R., Barr, C., & Stuart, R. (2000). Hedgerows as habitat for woodland plants. *Journal of environmental management, 60*(1), 77-90. doi: https://doi.org/10.1006/jema.2000.0363

Metzger, M. J., Brus, D. J., Bunce, R. G. H., Carey, P. D., Gonçalves, J., Honrado, J. P., Jongman, R. H. G., Trabucco, A., & Zomer, R. (2013). Environmental stratifications as the basis for national, European and global ecological monitoring. *Ecological Indicators, 33*, 26-35. doi: https://doi.org/10.1016/j.ecolind.2012.11.009

Michener, W., Brunt, J., & Vanderbilt, K. (2002). *Ecological informatics: A long-term ecological research perspective.* Paper presented at the Proceedings of the 6th World Multi-Conference on Systematics, Cybernetics and Informatics.

Michener, W. K. (2006). Meta-information concepts for ecological data management. *Ecological informatics, 1*(1), 3-7. doi: https://doi.org/10.1016/j.ecoinf.2005.08.004

Michener, W. K. (2015). Ecological data sharing. *Ecological informatics, 29*, 33-44. doi: https://doi.org/10.1016/j.ecoinf.2015.06.010

Michener, W. K., & Brunt, J. W. (2009). *Ecological data: design, management and processing*. Oxford: John Wiley & Sons.

Michener, W. K., Brunt, J. W., Helly, J. J., Kirchner, T. B., & Stafford, S. G. (1997). Nongeospatial metadata for the ecological sciences. *Ecological Applications, 7*(1), 330-342. doi: https://doi.org/10.1890/1051-0761(1997)007[0330:NMFTES]2.0.CO;2

Michener, W. K., & Jones, M. B. (2012). Ecoinformatics: supporting ecology as a data-intensive science. *Trends in ecology & evolution, 27*(2), 85-93. doi: https://doi.org/10.1016/j.tree.2011.11.016

Microsoft. (2020). Description of the database normalization basics Retrieved 22/09/20, 2020, from https://docs.microsoft.com/en-us/office/troubleshoot/access/database-normalization-description

Milner, C. (1975). Shetland project monitoring report Grange over Sands: Nature Conservancy Council.

Molloy, J. C. (2011). The open knowledge foundation: open data means better science. *PLoS biology, 9*(12). doi: https://doi.org/10.1371/journal.pbio.1001195

Moss, C. E. (1910). The fundamental units of vegetation: Historical development of the concepts of the plant association and the plant formation. *New Phytologist, 9*(1-2), 18-53.

National Geoscience Data Centre. (2020). National Geoscience Data Centre Retrieved 27/11/2020, from http://www.bgs.ac.uk/services/ngdc/

NERC. (2020). NERC Data policy Retrieved 27/11/2020, from https://nerc.ukri.org/research/sites/environmental-data-service-eds/policy/

Norton, L., Greene, S., Scholefield, P., & Dunbar, M. (2016). The importance of scale in the development of ecosystem service indicators? *Ecological Indicators, 61, Part 1*, 130-140. doi: https://doi.org/10.1016/j.ecolind.2015.08.051

Norton, L., Inwood, H., Crowe, A., & Baker, A. (2012a). Trialling a method to quantify the 'cultural services' of the English landscape using Countryside Survey data. *Land Use Policy, 29*(2), 449-455. doi: https://doi.org/10.1016/j.landusepol.2011.09.002

Norton, L. R., Maskell, L. C., Smart, S. M., Dunbar, M. J., Emmett, B. A., Carey, P. D., Williams, P., Crowe, A., Chandler, K., Scott, W. A., & Wood, C. M. (2012b). Measuring stock and change in the GB countryside for policy: key findings and developments from the Countryside Survey 2007 field survey. *Journal of Environmental Management, 113*, 117-127. doi: https://doi.org/10.1016/j.jenvman.2012.07.030

Open Data Institute. (2017). What is 'open data' and why should we care? Retrieved 3/12/2020, from https://theodi.org/article/what-is-open-data-and-why-should-we-care/

Peterken, G. F., & Backmeroff, C. (1988). *Long-term monitoring in unmanaged woodland nature reserves* (Vol. Issue 9). Peterborough: Nature Conservancy Council.

Petit, S., Griffiths, L., Smart, S. S., Smith, G. M., Stuart, R. C., & Wright, S. M. (2004). Effects of area and isolation of woodland patches on herbaceous plant species richness across Great Britain. *Landscape Ecology, 19*(5), 463-471. doi: https://doi.org/10.1023/B:LAND.0000036141.30359.53

Petit, S., Stuart, R., Gillespie, M., & Barr, C. (2003). Field boundaries in Great Britain: stock and change between 1984, 1990 and 1998. *Journal of Environmental Management, 67*(3), 229-238. doi: https://doi.org/10.1016/S0301-4797(02)00176-7

Pienta, A. M., & Lyle, J. (2018). Retirement in the 1950s. *IASSIST Quarterly, 42*(1), 12-12. doi: https://doi.org/10.29173/iq19

Pierik, M., Van Ruijven, J., Bezemer, T. M., Geerts, R. H., & Berendse, F. (2011). Recovery of plant species richness during long-term fertilization of a species-rich grassland. *Ecology, 92*(7), 1393-1398. doi: https://doi.org/10.1890/10-0210.1

Poore, M. (1955). The use of phytosociological methods in ecological investigations: I. The Braun-Blanquet system. *Journal of ecology, 43*(1), 226-244. doi: https://doi.org/10.2307/225713

Potter, C., & Lobley, M. (1996). The farm family life cycle, succession paths and environmental change in Britain's countryside. *Journal of Agricultural Economics, 47*(1-4), 172-190. doi: https://doi.org/10.1111/j.1477-9552.1996.tb00683.x

Rapp, A. (1996). Photo documentation of landscape change in northern Swedish mountains. *Ecological Bulletins*, 170-179. doi: https://www.jstor.org/stable/20113196

Ratcliffe, D. A. (1977). *A Nature Conservation Review: The Selection of Biological Sites of National Importance to Nature Conservation in Britain* (Vol. 2): Cambridge University Press.

Reichman, O. J., Jones, M. B., & Schildhauer, M. P. (2011). Challenges and opportunities of open data in ecology. *Science, 331*(6018), 703-705. doi: https://doi.org/10.1126/science.1197962

Rennie, S., Andrews, C., Atkinson, S., Beaumont, D., Benham, S., Bowmaker, V., Dick, J., Dodd, B., McKenna, C., & Pallett, D. (2020). The UK Environmental Change Network datasets–integrated and co-located data for long-term environmental research (1993–2015). *Earth System Science Data, 12*(1), 87-107. doi: https://doi.org/10.5194/essd-12-87-2020

Rhodes, C. J., Henrys, P., Siriwardena, G. M., Whittingham, M. J., & Norton, L. R. (2015). The relative value of field survey and remote sensing for biodiversity assessment. *Methods in Ecology and Evolution, 6*(7), 772-781. doi: https://doi.org/10.1111/2041-210X.12385

Ruttenberg, S., & Rishbeth, H. (1994). World data centres - past, present and future. *Journal of Atmospheric and Terrestrial Physics, 56*(7), 865-870. doi: https://doi.org/10.1016/0021-9169(94)90148-1

Sheail, J. (1987). Seventy-five years in ecology: the British Ecological Society Blackwell: Oxford.

Sheail, J., & Bunce, R. G. H. (2003). The development and scientific principles of an environmental classification for strategic ecological survey in the United Kingdom. *Environmental Conservation, 30*(2), 147-159. doi: https://doi.org/10.1017/S0376892903000134

Silvertown, J., Dodd, M. E., McConway, K., Potts, J., & Crawley, M. (1994). Rainfall, biomass variation, and community composition in the Park Grass Experiment. *Ecology, 75*(8), 2430-2437. doi: https://doi.org/10.2307/1940896

Silvertown, J., Poulton, P., Johnston, E., Edwards, G., Heard, M., & Biss, P. M. (2006). The Park Grass Experiment 1856–2006: its contribution to ecology. *Journal of Ecology, 94*(4), 801-814. doi: https://doi.org/10.1111/j.1365-2745.2006.01145.x

Smart, S., Ashmore, M., Hornung, M., Scott, W., Fowler, D., Dragosits, U., Howard, D., Sutton, M., & Famulari, D. (2005a). Detecting the signal of atmospheric N deposition in recent national-scale vegetation change across Britain. *Water, Air and Soil Pollution: Focus, 4*(6), 269-278. doi: https://doi.org/10.1007/s11267-005-3037-5

Smart, S., Bunce, R., Marrs, R., LeDuc, M., Firbank, L., Maskell, L., Scott, W., Thompson, K., & Walker, K. (2005b). 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 changes in trait representation. *Biological Conservation, 124*(3), 355-371. doi: https://doi.org/10.1016/j.biocon.2004.12.013
Smart, S., Maskell, L., Dunbar, M., Emmett, B., Marks, S., Norton, L., Rose, P., & Simpson, I. (2010a). An Integrated Assessment of Countryside Survey data to investigate Ecosystem Services in Great Britain. Countryside Survey Technical Report No. 10/07. Lancaster: NERC Centre for Ecology and Hydrology.
Smart, S. M., Bunce, R. G. H., Firbank, L. G., & Coward, P. (2002). Do field boundaries act as refugia for grassland plant species diversity in intensively managed agricultural landscapes in Britain? *Agriculture, Ecosystems and Environment, 91*, 73-87. doi: https://doi.org/10.1016/S0167-8809(01)00259-6
Smart, S. M., Clarke, R. T., van de Poll, H. M., Robertson, E. J., Shield, E. R., Bunce, R. G. H., & Maskell, L. C. (2003). National-scale vegetation change across Britain: an analysis of sample-based surveillance data from the Countryside Surveys of 1990 and 1998. *Journal of Environmental Management, 67*, 239-254. doi: https://doi.org/10.1016/S0301-4797(02)00177-9
Smart, S. M., Ellison, A. M., Bunce, R. G., Marrs, R. H., Kirby, K. J., Kimberley, A., Scott, A. W., & Foster, D. R. (2014). Quantifying the impact of an extreme climate event on species diversity in fragmented temperate forests: the effect of the October 1987 storm on British broadleaved woodlands. *Journal of Ecology, 102*(5), 1273-1287. doi: https://doi.org/10.1111/1365-2745.12291
Smart, S. M., Henrys, P. A., Purse, B. V., Murphy, J. M., Bailey, M. J., & Marrs, R. H. (2012). Clarity or confusion?: problems in attributing large-scale ecological changes to anthropogenic drivers. *Ecological Indicators, 20*, 51-56. doi: https://doi.org/10.1016/j.ecolind.2012.01.022
Smart, S. M., Marrs, R. H., Le Duc, M. G., Thompson, K., Bunce, R. G. H., Firbank, L. G., & Rossall, M. J. (2006). Spatial relationships between intensive land cover and residual plant species diversity in temperate farmed landscapes. *Journal of Applied Ecology, 43*, 1128-1137. doi: https://doi.org/10.1111/j.1365-2664.2006.01231.x
Smart, S. M., Scott, W. A., Whitaker, J., Hill, M. O., Roy, D. B., Critchley, C. N., Marina, L., Evans, C., Emmett, B. A., Rowe, E. C., Crowe, A., & Marrs, R. H. (2010b). Empirical realised niche models for British higher and lower plants – development and preliminary testing. *Journal of Vegetation Science, 21*, 643-656. doi: https://doi.org/10.1111/j.1654-1103.2010.01173.x
Southward, A. J., Langmead, O., Hardman-Mountford, N. J., Aiken, J., Boalch, G. T., Dando, P. R., Genner, M. J., Joint, I., Kendall, M. A., & Halliday, N. C. (2005). Long-term oceanographic and ecological research in the Western English Channel. *Advances in marine biology, 47*, 1. doi: https://doi.org/10.1016/S0065-2881(04)47001-1
Specht, A., Bolton, M. P., Kingsford, B., Specht, R. L., & Belbin, L. (2018). A story of data won, data lost and data re-found: the realities of ecological data preservation. *Biodiversity data journal*(6). doi: https://doi.org/10.3897/BDJ.6.e28073
SQL World. (2020). Database Normalization with Examples Retrieved 22/09/2020, 2020, from https://www.complexsql.com/database-normalization/
Stace, C. (1997). *New flora of the British Isles*. Cambridge: Cambridge University Press.
Ståhl, G., Allard, A., Esseen, P.-A., Glimskär, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Torell, Å. G., & Högström, M. (2011). National Inventory of Landscapes in Sweden (NILS)—scope, design, and experiences from establishing a

multiscale biodiversity monitoring system. *Environmental monitoring and assessment, 173*(1-4), 579-595. doi: https://doi.org/10.1007/s10661-010-1406-7
Steele, R. C. (1968). *National survey of woodlands. Reprinted from Society for the Promotion of Nature Reserves Handbook* (Vol. Reprinted from Society for the Promotion of Nature Reserves Handbook). Lincoln.
Steven, H. M., & Carlisle, A. (1959). *The native pinewoods of Scotland*. Edinburgh: Oliver & Boyd.
Stevens, C. J., Payne, R. J., Kimberley, A., & Smart, S. M. (2016). How will the semi-natural vegetation of the UK have changed by 2030 given likely changes in nitrogen deposition? *Environmental pollution, 208*, 879-889. doi: https://doi.org/10.1016/j.envpol.2015.09.013
Symstad, A. J., Chapin, F. S., Wall, D. H., Gross, K. L., Huenneke, L. F., Mittelbach, G. G., Peters, D. P., & Tilman, D. (2003). Long-term and large-scale perspectives on the relationship between biodiversity and ecosystem functioning. *Bioscience, 53*(1), 89-98. doi: https://doi.org/10.1641/0006-3568(2003)053[0089:LTALSP]2.0.CO;2
Tani, A., Candela, L., & Castelli, D. (2013). Dealing with metadata quality: The legacy of digital library efforts. *Information Processing & Management, 49*(6), 1194-1205. doi: https://doi.org/10.1016/j.ipm.2013.05.003
Tansley, A. (1939). The British Islands and their vegetation. Cambridge Cambridge University Press.
Tenopir, C., Rice, N. M., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLOS ONE, 15*(3), e0229003. doi: http://doi.org/10.1371/journal.pone.0229003
The Hedgerows Regulations. (1997). The Hedgerows Regulations (pp. pp16). London: The Stationery Office.
Turner, M. G., Collins, S. L., Lugo, A. L., Magnuson, J. J., Rupp, T. S., & Swanson, F. J. (2003). Disturbance dynamics and ecological response: the contribution of long-term ecological research. *BioScience, 53*(1), 46-56. doi: https://doi.org/10.1641/0006-3568(2003)053[0046:DDAERT]2.0.CO;2
UK Polar Data Centre. (2020). UK Polar Data Centre Retrieved 27/11/2020, from https://www.bas.ac.uk/data/uk-pdc/
UKCEH. (2020a). CAST (CEH Analytical Services Thesaurus) Retrieved 30/11/2020, from http://onto.nerc.ac.uk/CAST/0
UKCEH. (2020b). UK Centre for Ecology & Hydrology Retrieved 27/11/2020, from www.ceh.ac.uk
UKCEH. (2020c). UKCEH Land Cover Maps Retrieved 27/11/2020, from https://www.ceh.ac.uk/ukceh-land-cover-maps
UKRI-NERC. (2020). Data management planning Retrieved 30/11/2020, from https://nerc.ukri.org/research/sites/environmental-data-service-eds/dmp/
UKRI. (2020). Concordat on Open Research Data Retrieved 03/12/2020, from https://www.ukri.org/wp-content/uploads/2020/10/UKRI-020920-ConcordatonOpenResearchData.pdf
van den Berg, L. J., Jones, L., Sheppard, L. J., Smart, S. M., Bobbink, R., Dise, N. B., & Ashmore, M. R. (2016). Evidence for differential effects of reduced and oxidised nitrogen deposition on vegetation independent of nitrogen load. *Environmental Pollution, 208*, 890-897. doi: https://doi.org/10.1016/j.envpol.2015.09.017

Viking Energy. (2016). The Project (The Viking Wind Farm) Retrieved 7/1/2016, from http://www.vikingenergy.co.uk/the-project

Watson, R., Albon, S., Aspinall, R., Austen, M., Bardgett, B., Bateman, I., Berry, P., Bird, W., Bradbury, R., & Brown, C. (2011). *UK National Ecosystem Assessment: Technical report*. Cambridge: United Nations Environment Programme World Conservation Monitoring Centre.

Webb, D. A. (1954). Is the classification of plant communities either possible or desirable. *Bot. Tidsskr, 51*, 362-370.

Whitlock, M. C. (2011). Data archiving in ecology and evolution: best practices. *Trends in Ecology & Evolution, 26*(2), 61-65. doi: https://doi.org/10.1016/j.tree.2010.11.006

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin Core: an evolving community-developed biodiversity data standard. *PloS one, 7*(1), e29715. doi: https://doi.org/10.1371/journal.pone.0029715

Wilkinson, C., Brönnimann, S., Jourdain, S., Roucaute, E., Crouthamel, R., Brohan, P., Valente, A., Brugnara, Y., Brunet, M., & Team, I. (2019). Best Practice Guidelines for Climate Data Rescue: Copernicus Climate Change Service.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., 't Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., & Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*(1), 160018. doi: 10.1038/sdata.2016.18

Williams, W., & Lambert, J. M. (1959). Multivariate methods in plant ecology: I. Association-analysis in plant communities. *The Journal of Ecology, 47*(1), 83-101. doi: https://doi.org/10.2307/2257249

Williamson, F. (2016). Cross-disciplinary working in the sciences and humanities: Historical data rescue activities in Southeast Asia and beyond. *Geoscience Letters, 3*(1), 31. doi: https://doi.org/doi.org/10.1186/s40562-016-0062-3

Wood, C., Howard, D., Henrys, P., Bunce, R., & Smart, S. (2012). Countryside Survey: measuring habitat change over 30 years: 1978 data rescue-final report. Lancaster: Centre for Ecology & Hydrology.

Wood, C. M. (2019). National field survey takes advantage of innovative technology, 30/11/2020, from https://www.ceh.ac.uk/news-and-media/blogs/national-field-survey-takes-advantage-innovative-technology

Wood, C. M., Bunce, R. G., Norton, L. R., Smart, S. M., & Barr, C. J. (2018a). Land cover and vegetation data from an ecological survey of 'key habitat' landscapes in England, 1992-1993. *Earth System Science Data, 10*(2), 899-918. doi: https://doi.org/10.5194/essd-10-899-2018

Wood, C. M., & Bunce, R. G. H. (2016a). Ecological survey of the native pinewoods of Scotland 1971. *Earth Syst. Sci. Data, 8*(1), 177-189. doi: https://doi.org/10.5194/essd-8-177-2016

Wood, C. M., & Bunce, R. G. H. (2016b). Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring. *Earth Syst. Sci. Data, 8*(1), 89-103. doi: https://doi.org/10.5194/essd-8-89-2016

Wood, C. M., Bunce, R. G. H., Norton, L. R., Maskell, L. C., Smart, S. M., Scott, W. A., Henrys, P. A., Howard, D. C., Wright, S. M., Brown, M. J., Scott, R. J., Stuart, R. C., & Watkins, J. W. (2018b). Ecological landscape elements: long-term monitoring in Great Britain, the Countryside Survey 1978–2007 and beyond. *Earth Syst. Sci. Data, 10*(2), 745-763. doi: https://doi.org/10.5194/essd-10-745-2018

Wood, C. M., Smart, S. M., & Bunce, R. G. H. (2015). Woodland Survey of Great Britain 1971–2001. *Earth Syst. Sci. Data, 7*(2), 203-214. doi: https://doi.org/10.5194/essd-7-203-2015

Wood, C. M., Smart, S. M., Bunce, R. G. H., Norton, L. R., Maskell, L. C., Howard, D. C., Scott, W. A., & Henrys, P. A. (2017). Long-term vegetation monitoring in Great Britain – the Countryside Survey 1978–2007 and beyond. *Earth Syst. Sci. Data, 9*(2), 445-459. doi: https://doi.org/10.5194/essd-9-445-2017

Worthington, E. (1965). The international biological programme. *Nature, 208*(5007), 223-226. doi: https://doi.org/10.1038/208223a0

Zimmerman, A. S. (2008). New knowledge from old data: The role of standards in the sharing and reuse of ecological data. *Science, Technology, & Human Values, 33*(5), 631-652. doi: https://doi.org/10.1177/0162243907306704

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of e-Government, 10*(2).

Earth Syst. Sci. Data, 7, 203–214, 2015
www.earth-syst-sci-data.net/7/203/2015/
doi:10.5194/essd-7-203-2015

Open Access Earth System Science Data

Woodland Survey of Great Britain 1971–2001

C. M. Wood[1], S. M. Smart[1], and R. G. H. Bunce[2]

[1]Centre for Ecology and Hydrology, Lancaster Environment Centre, Bailrigg, Lancaster, LA1 4AP, UK
[2]Estonian University of Life Sciences, Kreuzwaldi 5, 51014 Tartu, Estonia

Correspondence to: C. M. Wood (clamw@ceh.ac.uk)

Received: 10 February 2015 – Published in Earth Syst. Sci. Data Discuss.: 25 February 2015
Revised: 13 July 2015 – Accepted: 14 July 2015 – Published: 10 August 2015

Abstract. The Woodland Survey of Great Britain is a unique data set, consisting of a detailed range of ecological measurements at a national scale, covering a time span of 30 years. A set of 103 woods spread across Britain were first surveyed in 1971, which were again surveyed in 2000–2003 (for convenience referred to subsequently as the "2001 survey"). Standardised methods of describing the trees, shrubs, ground flora, soils and general habitats present were used for both sets of surveys. The sample of 1648 plots spread through 103 woodland sites located across Britain makes it probably the most extensive quantitative ecological woodland survey undertaken in Britain; it is also notable for the range of sites that have been revisited after such a long interval. The data set provides a unique opportunity to explore the effects of a range of potential drivers of woodland change that operated between 1971 and 2001. The data set is available in four discrete parts, which have been assigned the following DOIs: doi:10.5285/4d93f9ac-68e3-49cf-8a41-4d02a7ead81a (Kirby et al., 2013b), doi:10.5285/d6409d40-58fe-4fa7-b7c8-71a105b965b4 (Kirby et al., 2013d), doi:10.5285/fb1e474d-456b-42a9-9a10-a02c35af10d2 (Kirby et al., 2013c), doi:10.5285/2d023ce9-6dbe-4b4f-a0cd-34768e1455ae (Kirby et al., 2013a).

1 Introduction

In 1971, a national survey of semi-natural woodlands in Great Britain was undertaken at the Nature Conservancy's research station at Merlewood, Grange over Sands, Cumbria (a predecessor of the Centre for Ecology and Hydrology). The survey of 103 sites was planned by R. G. H. Bunce and M. W. Shaw (Bunce and Shaw, 1972; Hill et al., 1975; Bunce, 1981). The project at this time had the following objectives:

1. To develop an efficient user-orientated method of classifying semi-natural woodland ecosystems in Britain.

2. To develop a complementary method of phytosociological classification for semi-natural woodlands.

3. To use or assist in the use of the classification in the fulfilment of the Nature Conservancy's aims and polices for wildlife conservation (Bunce and Shaw, 1973a).

Within the 103 woodland sites chosen, ecological information was recorded at the site level and in more detail from 16 200 m^2 sample plots located at random within each site. From each of these plots the following data were collected: presence of vascular plants and bryophytes from five nested quadrat sizes, measurement of diameters at 1.3 m (DBH – diameter at breast height) of all trees over 5cm in diameter in the plot and of saplings and shrubs in specified quarters of the plot, site descriptions and soil samples. These data were collected from the 103 sites (1648 plots) by eight survey teams between July and September 1971.

In 2000, it was thought timely to revisit the 1971 survey. This time, the survey was focused on assessing the changes that had occurred within the woodland sites in the intervening 30 years, moving away from the original goals of the 1971 survey as outlined above. Fourteen sites were visited in 2000 as part of a pilot survey to assess the logistical and analytical implications of trying to carry out a re-survey (Smart et al., 2001). No surveys were carried out in 2001 because of a serious foot-and-mouth disease outbreak in livestock (during which access to the British countryside was severely restricted in order to constrain the contagious disease) but

56 sites were surveyed in summer 2002 and the remainder in 2003 by teams of consultant ecologists using exactly the same field methods as in 1971, as described below. Prior to each survey, a two-day training course was held at the Centre for Ecology and Hydrology to thoroughly prepare the surveyors with the detailed field protocols. Additionally, in 1971, all survey teams were initially accompanied by a supervisor and regular visits to the field were made by the project leader to ensure consistency and quality in data recording according to criteria laid out in the field handbook (Shaw and Bunce, 1971). In the 2001 survey, experienced survey staff were available in the office to answer post-training queries from the field throughout the survey via telephone and a full quality assurance exercise was carried out as described below, and more fully in Kirby et al. (2005).

2 Survey sites

The 103 surveyed woodlands were chosen from a set of 2453 woodlands that had been part of a preliminary survey known as the "Steele" survey (Steele, 1968). This had begun in the late 1960s and was led by R. C. Steele, the head of the Nature Conservancy's Woodland Management section. Standard recording cards were used, and the data provided background information for the Nature Conservation Review (Ratcliffe, 1977).

The subset of 103 was derived from the 2453 by association analysis (Williams and Lambert, 1959) and other numerical techniques that, at the time, were still novel and undergoing rapid development (Hill et al., 1975; Bunce, 1981; Bunce and Shaw, 1973b). These analyses put the woods into 103 groups according to the similarity of their plant species composition. The wood that was most typical of that group was then selected for detailed survey. Site names and grid references are given in Table 1 (it should be noted that the majority of the sites are in private ownership and therefore permission from the landowner must be sought before any potential visit).

2.1 Site descriptions

The sites provide a representative sample of the geographic spread of woodland cover (see Fig. 1) and the range of broadleaved/semi-natural woodland types. The sites also show a considerable physiographic variability in terms of rainfall, slope and aspect (Corney et al., 2004). The number of sites recorded in the 1971 survey from each of the 32 original ITE land classes in Britain (Bunce et al., 1990) was compared with the mean percentage area of broadleaved woodland, estimated from Countryside Survey 2000 data (Haines-Young et al., 2000), for each land class (Bunce et al., 1996). The comparison shows a good correspondence between national woodland area and the number of woodland survey sites recorded, with proportionally more surveyed woods

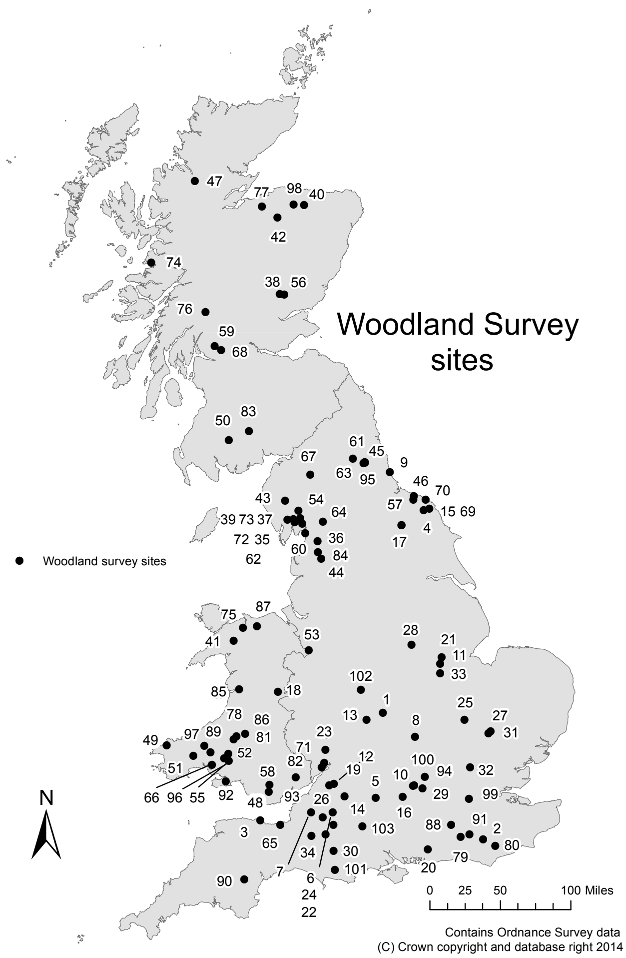


Figure 1. Map of woodland survey site locations across Great Britain.

from land classes with a high broadleaved woodland cover (Kirby et al., 2005) (see Table 2).

Additionally, we can compare the number of plots allocated to each National Vegetation Classification (NVC) group (Rodwell, 1991) with the estimated total area of NVC types in ancient semi-natural woodland across Britain (Cooke and Kirby, 1994) (see Table 3). The 1971 survey data span the broad range of types in roughly the proportions that might be expected from the Cooke and Kirby data. Secondly, a comparison was made with the sample of woody vegetation from the GB Countryside Survey from 2000 (Haines-Young et al., 2000). The 1971 plots were grouped by Countryside Vegetation System classes (Bunce et al., 1999) and their frequency was compared to the estimated national area of each class. The two data sets are generally well correlated (Kirby et al., 2005).

In terms of woodland size, the woods surveyed range from 4 to 100 ha with a single outlier of 312 ha (Glen Beasdale Wood, Scotland). The mean size of the sample was 31.8 ha and the median 20.4 ha. The lower size cut-off was deter-

Table 1. List of the 103 woodland sites.

Site code	Site	OSGB easting	OSGB northing	Site code	Site	OSGB easting	OSGB northing
1	Waverley Wood	4355	2710	53	Bubney Wood	3509	3420
2	Pickreed Wood	5503	1266	54	Newclose Wood	3392	5015
3	Greenaleigh Plantation	2955	1479	55	Carmel Wood	2594	2162
4	Reins Wood	4567	4850	56	Den of Alyth Wood	3230	7487
5	Love's Copse	4274	1735	57	Pinkney Bank Wood	4704	5142
6	Longleat Woods	3790	1432	58	Coed Gelli-draws	3058	1885
7	Compton Wood	3537	1570	59	Gartfairn Wood	2434	6896
8	Say's Copse & Smalladine Copse	4724	2435	60	Eaves Wood	3468	4762
9	Hawthorn Dene	4435	5458	61	Longclose Wood	4135	5560
10	Kitesgrove, Juniper Hall and Home woods; Big Ashes and Stockings Plantations	4715	1880	62	Winster Wood	3410	4930
11	Old Park Wood	5011	3267	63	Riding Mill Wood	4013	5612
12	Midger Wood & Back Common	3797	1895	64	Rottenbutts Wood	3670	4890
13	Austy Wood	4170	2627	65	Great Plantation	3183	1431
14	Birds Marsh	3918	1756	66	Glan Morlais	2403	2114
15	Beck Hole Scar	4823	5022	67	Eden Gorge Wood	3527	5425
16	Ashampstead Common	4582	1750	68	Blane Wood	2507	6851
17	Ashberry Wood	4569	4851	69	Newton House Wood	4885	5040
18	Ffridd Wood	3157	2947	70	Over Dale Wood	4847	5140
19	Lower Wetmoor	3742	1877	71	Morse's Grove	3685	2137
20	Wellhanger Copse	4870	1147	72	Hall Brow	3348	4885
21	Sapperton South Wood & Pickworth Wood	5030	3340	73	Great Knott	3334	4918
22	Park Wood	3703	1321	74	Glen Beasdale Wood	1708	7847
23	Betty Daw's Wood	3698	2283	75	Ceunant Dulyn	2757	3683
24	Hill Wood	3782	1574	76	Coille Coire Chuilc	2327	7281
25	Papworth Wood	5291	2629	77	Dounduff Wood	2975	8486
26	Loocombe Wood	3668	1512	78	Allt-yr-Hebog	2685	2440
27	Rivey Wood	5565	2478	79	Warren Wood	5245	1294
28	Spital	4683	3484	80	Hoad's Wood	5643	1187
29	Medmenham Wood	4810	1845	81	Wern-fawr Wood	2588	2239
30	Piddles Wood	3795	1130	82	Blakeneyhill Wood	3658	2087
31	Balsham	5588	2496	83	Tynron	2825	5924
32	Hoddesdonpark Wood	5353	2085	84	Wellington Wood	3513	4546
33	Docksight Wood	5013	3158	85	Allt-ddu and Dol-y-garnedd Wood	2715	2973
34	Luns Hill Wood	3539	1307	86	Dinas Wood	2783	2467
35	Whitbarrow Wood	3436	4870	87	Coedcochion Wood	2916	3694
36	Pike Gill Wood	3610	4668	88	Leith Hill Place Wood; Farmhouse, Slittens & Hooks Copses	5137	1427
37	Birks Brow	3410	4920	89	Allt Blaen-eigiau	2384	2256
38	Craighall Gorge	3178	7490	90	Houndtor Wood	2770	804
39	Haverigg Holme	3264	4915	91	Chiddingly Wood	5347	1320
40	Mill Wood	3455	8504	92	Gelli-hir Wood	2563	1927
41	Coed Y Wenault	2649	3531	93	Llangibby Park Wood	3360	1972
42	Callender	3150	8367	94	Bradenham Wood; The Coppice	4835	1975
43	Seatoller Wood	3239	5131	95	Priestfield	4153	5568
44	New Laund and High Wood	3653	4468	96	Garreg-goch-isaf Wood	2540	2185
45	Sliding Braes	4148	5569	97	Afon Sylgen Wood	2315	2332
46	White Cliff Wood	4711	5185	98	Glen Orchill Wood	3335	8516
47	Corrieshalloch Gorge	2205	8780	99	Dulwich Wood	5340	1725
48	Hensol Wood	3052	1802	100	Nettlebed Common Wood	4700	1875
49	Pen-yr-allt Wood	1884	2338	101	Oakers	3808	916
50	Garroch Wood	2595	5822	102	Lower Nut Hurst Wood	4105	2970
51	Cil-Hen-Ros	2188	2215	103	Normanton Down Gorse	4121	1414
52	Allt Penarth Wood	2648	2407				

Table 2. Comparisons of numbers of sites surveyed (1971) and broadleaved woodland area (Countryside Survey 2000 data) for different environmental zones (aggregations of land classes) (Carey et al., 2008).

Environmental Zone	Number of woods surveyed	Broadleaved woodland area (thousand ha)
Easterly lowlands (England)	33 (32 %)	489 (37 %)
Westerly lowlands (England)	31 (30 %)	400 (30 %)
Uplands (England)	5 (5 %)	38 (3 %)
Lowlands (Scotland)	6 (6 %)	118 (9 %)
Intermediate uplands and islands (Scotland)	5 (5 %)	52 (4 %)
True uplands (Scotland)	2 (2 %)	59 (4 %)
Lowlands (Wales)	15 (15 %)	75 (6 %)
Uplands (Wales)	6 (6 %)	97 (7 %)
Total	103 (100 %)	1328 (100 %)

Table 3. Comparison of woodland NVC types identified in 1971 survey data with other estimates across the country. Values expressed as percentages of totals.

NVC grouping	Number of NVC samples Rodwell (1991)	1971 records	Area estimates from Cooke and Kirby (1994)
Mesotrophic oakwoods (W10, 11)	27	53	42
Ash–elm woods (W8, 9)	28	18	29
Acidic oakwoods (W16, 17)	9	12	14
Alderwoods (W5–7)	14	6	6
Birch–willow woods (W1–4)	9	2	5
Calcareous beech–yew woods (W2, 13)	7	4	2
Acidic beechwoods (W14, 15)	6	5	2

mined by the minimum size (10 acres/4 ha) used in the original "Steele" survey (Steele, 1968).

2.2 Plot layout and descriptions

Sixteen plots were randomly positioned within each site in 1971 and the location of each was marked on a 1 : 25 000 map. Each plot was 14.1×14.1 m (200 m^2) (Fig. 2) and constructed as shown in Fig. 3, with one centre post and four corner posts, with a set of four strings tagged with markers at specified distances. The centre post had a right-angled gauge affixed to the top in order to orientate the plot at random. In the field, plots were located by pacing from the nearest relocatable feature. Data were then collected on ground flora, tree and shrub layers, soils and habitat characteristics for the plot as described below. A habitat sheet for the whole wood was also compiled.

In the 2001 survey, the original maps were used to find the same plot position from 1971 as accurately as possible. Analysis of the 1648 plot records taken in 1971 and 2001 described in Kirby et al. (2005) demonstrates that the records may be treated as paired data (i.e. relocation error was not significant, as described in the "Data quality" section below). The advantage of paired data is that derived variables, such as species richness, can be reduced to differences for purposes of statistical testing. The total variation across time and sites will be less than if two completely random samples were collected in each year and the power of tests is thereby increased. Some relocation error was, however, inevitable given the limited information available.

2.3 Methodology in context

It is often an insoluble problem that, in order to extend an older time series without breaking consistency with its established methods, methods have to be repeated despite a more modern design perhaps being preferable if we were to start again. However, although the protocols in question are old, that does not necessarily mean they are outdated.

This survey was the first time at a national level that samples were being used to obtain an integrated assessment of the response of vegetation to the environment across a defined population. The structure of the project provided the basis for the further development of strategic survey methods. A subsequent survey based on these methods (Bunce and Shaw, 1973b), the Classification of the Native Pinewoods of Scotland, set the conservation agenda for that scarce resource (Bunce and Jeffers, 1977). In later work, the concept of a woodland site, and subsequently a 1 km square sampled at random, with random plots sampled within, be-

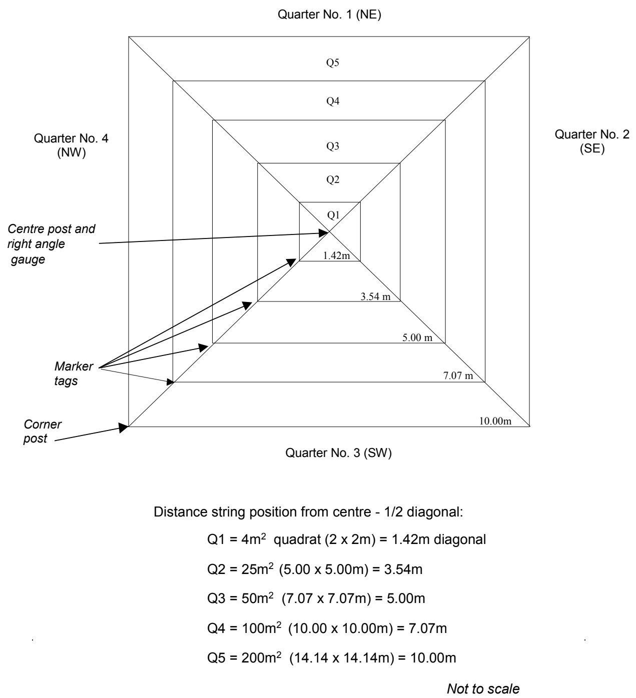


Figure 2. Layout of vegetation plot.

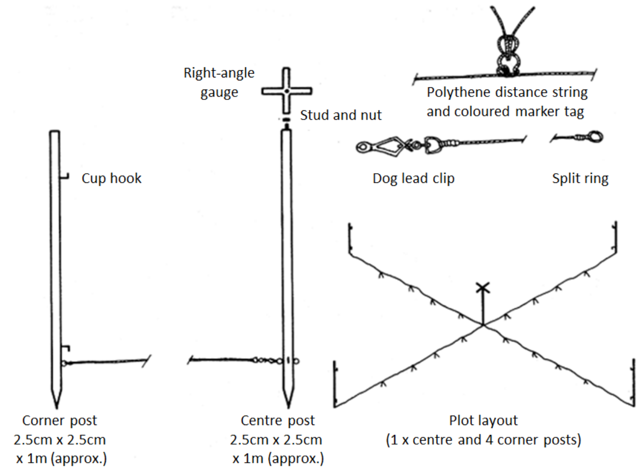


Figure 3. Plot construction.

came a standard sampling strategy used as the basis of surveys such as the Cumbria Survey (Bunce and Smith, 1978) and the Terrestrial Survey of Shetland (Milner, 1975). Variations of this method are currently used very successfully in several other large ecological surveys in Britain, such as the Countryside Survey (Carey et al., 2008) and the Glastir Monitoring and Evaluation Programme (Emmett and GMEP team, 2014). Within the European Biodiversity Observation Network (EBONE), methods adapted from the basic principles in this woodland survey have been developed to roll out across the whole of Europe (Bunce et al., 2008, 2011). During the EBONE project, the methods were widely tested across 12 European countries, and also Israel, Australia and South Africa. The methods were proven to be robust, reliable and repeatable at a continental, landscape scale (Roche and Geijzendorffer, 2013).

A key aim of the sampling design was that the methods chosen should be standardised, and therefore highly repeatable. The size of the plot was chosen with reference to continental phytosociologists who at the time most widely used plots of between 100 and 200 m^2 (Bunce and Shaw, 1973b). After preliminary field tests, it was found that the number of species recorded usually stabilised at this size. The area of 200 m^2 was thus adopted for this survey, with five nests within. As the focus of the survey is on ground flora as well as tree and shrub information, the square plot with inner nests aids a systematic search of the vegetation within the plot. It is also straightforward to lay out in the field, and ensures a standard-sized plot is laid out every time. For these reasons, we consider the square plot as more advantageous than a circular plot. Plotless sampling was also dismissed, as it is not a suitable method for recording ground vegetation, only tree density. Random sampling was preferred to systematic sampling in this case to avoid the possibility of resonance with environmental features, for example a map grid line following the course of a stream. Random sampling also has practical advantages over systematic sampling, which requires continuous scale adjustment in order to obtain a constant sample from variably sized areas (Bunce and Shaw, 1973b).

3 Data collected

3.1 Site information, plot locations and information, slope and aspect

For both the whole woodland site, and for each of the 16 200 m^2 plots within, the presence and absence of a series of attributes were recorded. Attributes included management factors such as the presence of coppice or stumps, physiological factors such as the presence of rock or cliffs, habitat-related factors such as the presence of rotting stumps or hollow trunks, aquatic habitats such as ponds, presence of buildings or open habitats such as glades and rides, presence of epiphytes on trees, presence of animals and birds, and also boundary types and nearby land use. A full list of habitats may be found in the 1971 field handbook (Shaw and Bunce, 1971) (supplied as supporting documentation with the data sets). The slope of each plot was measured in degrees using a hypsometer and the aspect of each plot was measured using a magnetic compass.

3.2 Vegetation data

Within the plot described in Fig. 2, the area within the first nest of the plot (2 × 2 m) was searched for the presence of all vascular plants (monocots, dicots and ferns), including tree species. This procedure was repeated for each nest of the quadrat, increasing the size each time as shown in Fig. 2. In the final nest (the whole 200 m^2 plot), the percentage cover (to the nearest 5 %) of each species was estimated. In addition, the total cover of bryophytes was estimated from the entire plot, as was an overall estimate for litter, wood, rock, bare ground and standing water. Bryophytes and lichens were collected separately and specimens identified later in 1971; in the 2001 survey only a limited list of common bryophytes was recorded. Some species were recorded in 1971 as amalgamated taxa reflecting difficulties in their consistent separation, for example *Quercus robur* and *Q. petraea*. In the data set, amalgamated taxon codes have been applied in order to remove the effect of recorders separating out such species to differing degrees.

3.3 Soil data

In both 1971 and 2001, soil samples were taken from every accessible plot in every woodland. A single composite soil sample was taken from each plot, at the centre of the vegetation quadrat, using a trowel. Samples (weighing approximately 1 kg) were taken to a depth of 15 cm and placed in a labelled plastic bag. On return to the laboratory, all soil samples were stored at 4 °C prior to processing and analyses. Soil samples from the 2001 survey were sieved using a 2 mm automatic sieving machine. A pH reading was taken on a representative fresh subsample from each soil sample before air-drying at 20 °C. Another subsample was then taken to determine loss on ignition (LOI), as a measure of soil organic matter content. Unless otherwise stated, soil pH values in the data set are from the soil samples prior to air-drying ("fresh").

All analyses were carried out under the supervision of the Environmental Chemistry Section at the Centre for Ecology and Hydrology (CEH), Merlewood, following standard methodologies and quality control procedures (Allen, 1989), including the analyses of certified standard reference samples within batches.

During the 2001 survey, the same soil analysis protocols were used as in the 1971 survey but the equipment was different. Changes in analytical precision since 1971, due to modifications in technical equipment, could have influenced the significance of the results obtained from both pH and LOI. Therefore repeat analyses of LOI on the 1971 samples and comparisons between fresh and air-dried soil samples from 1971 and 2001 were done to check the comparability of analytical methods between the two surveys. A representative number (ca. 20 %) of soil samples from 1971 were analysed for pH and LOI using the same procedures and equipment as for the 2001 survey. These results are included in the published data set.

Soil group information is derived from data recorded in 1971. Information on soil moisture, texture, structure and colour for different horizons was recorded in the field. This information was translated into comparable Avery (1980) soil codes in 2001.

3.4 Tree diameter

Trees, saplings and shrubs were recorded in the 200 m^2 plot, as described above. Decisions as to whether individuals are in the plot or not were based on the rooted base being 50 % or more within the plot.

For trees (stems of more than 5 cm diameter at breast height (DBH) of any species normally capable of attaining a treelike habit in Britain), the species and DBH of all stems in the whole plot greater than 5 cm were measured. Trees with multiple stems had each stem recorded separately. Standing dead trees were also measured and identified as such.

Saplings (definition as for trees, but with a height of less than 130 cm and with a DBH less than 5 cm) were recorded only in quarters 1 and 3 of the plot (see Fig. 2). The same measurements as for trees were made. Shrubs, like saplings, were also only recorded in quarters 1 and 3, and again the same measurements were taken. Shrubs were defined as species including hazel, blackthorn, *Viburnum* spp. and juniper. See Table 4 for a summary of data collected.

4 Data quality

The 1971 data sets were transferred from the original field sheets to spreadsheets prior to the 2001 surveys. The 1971 data were double-punched and then checked and corrected to produce a final validated copy. In the 2001 surveys, the consultant surveyors were asked to ensure that all data were corrected and validated prior to transfer in electronic form to CEH. Initial standard validation checks included plot and site counts to ensure no duplicate numbering and hence double counting of plots.

As part of the quality assurance process for the ground flora data, six sites were visited by a different set of surveyors and eight plots at each site recorded within 2 weeks of the main survey. A mixed model ANOVA showed no overall difference in species richness between the different surveyors (Kirby et al., 2005).

Some plot relocation error was inevitable given the limited information available and the nature of the original maps. In the repeat survey, the field botanist relied only on a marked point on a map as the sole aid to relocating the 1971 plot location. As statistical analyses of temporal vegetation change are more powerful when based on records from plots located in the same place rather than randomised to new locations at each survey, a method was developed to measure whether the 2001 record for a plot was more similar to the record for that plot in 1971 than another (randomly chosen) position from 1971. This follows from the general principle that locations near to each other tend to be more similar. Therefore, the principle of autocorrelation between near points was used to

Table 4. Summary of data collected.

Data category	1971 survey	2001 survey
Ground flora	Species present in the plot % cover/ abundance estimates bryophyte collection	As 1971 except that only most common bryophytes recorded.
Trees	DBH (diameter at breast height) and species recorded from all four quadrants	As 1971
Shrubs and saplings	DBH and species recorded from diagonally opposite quarters	As 1971
Seedlings	Included with the ground flora records	As 1971
Plot description and habitats	Tick list of features (broad categories): Tree management Regeneration Dead trees Epiphytes Rock habitats Aquatic habitats Open habitats Human elements Vegetation structure Animal signs	As 1971
Soil data	Tick list description from small pit and augur boring in the centre of the plot – to determine soil type Composite soil sample from top 10–15 cm.	Composite soil sample from top 10–15 cm.
Whole wood description	Tick list of features (broad categories). As for plot, plus adjacent land use and boundary type	As 1971, plus surveyors were asked to make a summary report for the whole site (Site Surveyors, 2003).

address the problem of quantifying the error involved in attempting to relocate the same vegetation monitoring plots.

In attempting to measure the amount of relocation error, one cannot of course exploit a "true" set of temporal pairs known to have been recorded in exactly the same position. What can be done is to compare the average species compositional similarity between the ostensibly true temporal pairs with the average similarity for a random pairing of the 1971 data with the 2001 data. If, on average, attempts to relocate the true 1971 position had been successful then the similarity between the true pairs should be greater than the random pairs. This approach was tested on the 14 pilot resurvey sites (Smart et al., 2001). All the sites showed higher similarity between plots as a result of the search for the 1971 plot location, and for nine sites there was significantly higher similarity. The same analysis was carried out for all the remaining sites. Overall at 97 sites (out of 103) mean similarity was greater between "relocated" plot pairs compared to random-pair comparison; for 59 sites the difference was significantly greater. The data have therefore been improved through the identification of the original plot locations. There is still a need for caution in interpreting the explanatory power of plot-level variables because of the possible confounding of plot relocation error and change over time. Small differences between years in plot location, for example, from an open patch to a more shaded patch could result in lower species richness and higher woody basal area being recorded for that plot. However, given the size of the data set, individual plot errors due to this factor are likely to be balanced out over the whole sample. A full account of this is given in Appendix 3 of Kirby et al. (2005).

It is important to note that there were some marked differences in the date of surveys between 1971 and the 2001 surveys, with most sites being recorded earlier in the year in 2001. This is likely to influence the recorded presence or abundance of vernal species in particular, with more species generally detectable in the late April–July period (Kirby et al., 1986; Sykes et al., 1983; Sykes and Horrill, 1979) than much later surveys. More species records would therefore be expected from the 2001 surveys.

In terms of the analytical soil data, quality control measures were followed as outlined in Allen (1989). These included the analyses of certified standard reference samples within batches. The descriptive profile data collected in 1971 were collected following the standards set out in the training and field handbook but were not formally checked for quality aside from checks from supervisors during the survey.

5 The Woodland Survey in context

Although there are many schemes across the world that monitor trees and forestry, there are few long-term programmes that take an integrated approach such as the survey in question, including trees, but also vegetation and soil information. Many national forest and woodland monitoring schemes were initially set up with an emphasis on monitoring timber production, commonly in the 1920s, when timber supplies were low following the First World War. For example, in Britain, the Forestry Commission was set up in 1919; since then it has undertaken national forestry inventory surveys which concentrate on the size, distribution, composition and condition of all forests in Britain but does not focus on sampling ground flora or soils (Forestry Commission, 1952, 1970, 1984, 2003, 2013). The situation is similar in the heavily forested countries of northern Europe such as Sweden, Denmark and Finland, where national forest inventories are also carried out (Groom and Reed, 2001), and also in the United States of America, where the US Forest Service has had a monitoring programme in place since the 1920s (Smith, 2002; United States Forest Service, 2015a).

An additional driver for the initiation of forest surveys across central Europe was the mystery of *Waldsterben* (forest decline). This became a contentious issue in the early 1980s, when it was suggested that air pollution was causing a progressive death of forests (Hinrichsen, 1987). In Germany, the forest authorities initiated surveys of the national forests, starting in 1987 and repeated at approximately decadal intervals, and currently carried out by the Thünen Institute of Forest Ecosystems (Kändler, 2009; Kandler and Innes, 1995; Thünen Institute, 2015). In Switzerland also, a thorough national forest inventory was first carried out in the early 1980s, repeated in the mid-1990s and again in the mid-2000s. Since 2009, the inventory has become a continuous monitoring programme. The inventory records the current state and changes of the Swiss forest (Mandallaz, 2007; National Forest Inventory, 2015; Böhl and Brändli, 2007). In both of these countries, the inventories are, again, largely focused on monitoring timber production, although both have been concerned with forest condition from the start, and the Swiss inventory in particular has come to include greater detail regarding a range of habitat measures (as described in the field manuals, e.g. Keller, 2011).

In tropical regions there is a general shortage of biodiversity data (Balmford et al., 2005), which is largely due to the geographical inaccessibility of many of the areas, and lack of local resource. Many studies regarding forestry and woodland in these regions rely heavily on remotely sensed information and concentrate on extent, biomass and carbon stocks (Asner, 2015; DeVries et al., 2015; Sousa et al., 2015; Wani et al., 2015), rather than ground-level biodiversity at a national level. Efforts are being made in many countries to intensify soil and ground vegetation sampling, as in the USA (United States Forest Service, 2015b; Smith, 2002); however, it is important to remember that the focus of this British woodland survey is on the semi-natural woodland ecosystem (not only trees and shrubs but also soils and ground flora). Taking this into account, there is relatively little literature regarding comparable national long-term monitoring schemes across the world, particularly those dating back as far as 1971.

Examples of data usage

Since the first survey in 1971, the data have been analysed and used in a range of ways to answer a variety of questions. After the first survey in 1971, publications arising from the data included the production of "A Field Key for Classifying British Woodland Vegetation" (Bunce, 1982, 1989). The survey was also described in Bunce and Shaw (1972) and used to put British woodlands into a European context in Bunce (1981). The standardised methods, as described in Bunce and Shaw (1973b), became the basis for a range of subsequent large surveys, as described in Sect. 2.3.

Following the second survey in 2001, a range of analyses were undertaken, as described in Kirby et al. (2005), focusing on changes that had taken place between the two surveys. Some of the conclusions from the main findings were that there had been an overall increase in soil pH, particularly in organic soils, but there was no increase in the mean level of soil organic matter. Most tree and shrub species remained stable in terms of their frequency of occurrence at plot and site levels, although 15 species (9 of these shrubs) declined, whilst 5 other species (4 conifers) increased. There was a net loss of stems from the smallest size classes (particularly less than 10 cm DBH) with some smaller gains in the 30–60 cm classes. Stems greater than 60 cm remained scarce, although different species revealed distinct patterns of variation. Overall ground flora species richness declined by up to 32 % at a plot level (Kirby et al., 2005).

More recently, further studies have included an analysis of the impact of an extreme weather event – a storm in 1987 during which wind speeds locally gusted up to 160 kph and an estimated 15 million trees were blown down across the south of England. Using Bayesian methods, Smart et al. (2014) demonstrated that woodland plots inside the storm track had a lower loss of understorey species richness, or an increase in richness between 1971 and 2001.

Marrs et al. (2013) analysed the data in order to investigate the impact of aggressive dominant native species on the species richness of native woodlands. Findings suggested that several species do have the potential to become "overdominant" and perhaps may impinge on other field-layer species.

Table 5. Summary of data sets available.

Data set	DOI	Description
Site information		
Woodlands_Survey_Site_Information_1971_2001	doi:10.5285/d6409d40-58fe-4fa7-b7c8-71a105b965b4	Slope, aspect, locations, descriptions, habitat categories
Vegetation data		
Woodlands_Survey_Flora_Data_1971_2001	doi:10.5285/2d023ce9-6dbe-4b4f-a0cd-34768e1455ae	Vascular plants, bryophytes, lichens
Soil data		
Woodlands_Survey_Soil_Data_1971_2001	doi:10.5285/fb1e474d-456b-42a9-9a10-a02c35af10d2	pH and loss on ignition (soil organic matter), soil group
Tree diameter		
Woodlands_Survey_Tree_Diameter_Data_1971_2001	doi:10.5285/4d93f9ac-68e3-49cf-8a41-4d02a7ead81a	Trees – diameter at breast height (DBH)

Corney et al. (2006) undertook a multivariate analysis to assess the effects of landscape-scale environmental drivers on the vegetation composition of British woodlands. The analysis investigated the degree to which field-layer vegetation composition in forests is determined by variables operating at different scales, from regional (such as climate, location) to local factors (such as the basal area of canopy trees and management).

Additionally, the plot species data have contributed to Great Britain niche models such as MutiMOVE (Henrys et al., 2015). MultiMOVE is a statistical package that contains fitted niche models for almost 1500 plant species in Great Britain. The models have been fitted using multiple statistical techniques in order to make predictions of species occurrence from specified environmental data, including this woodland data. It also allows plotting of relationships between species' occurrence and individual covariates so that the user can see what effect each environmental variable has on the specific species in question.

6 Data availability

The data sets have been assigned digital object identifiers, and users of the data must reference the data as follows:

- Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., and Shaw, M. W.: Woodlands Survey Tree Diameter Data 1971–2001, NERC Environmental Information Data Centre, doi:10.5285/4d93f9ac-68e3-49cf-8a41-4d02a7ead81a, 2013.

- Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., and Shaw, M. W.: Woodlands Survey Site Information 1971–2001, NERC Environmental Information Data Centre, doi:10.5285/d6409d40-58fe-4fa7-b7c8-71a105b965b4, 2013.

- Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., and Shaw, M. W.: Woodlands Survey Soil Data 1971–2001, NERC Environmental Information Data Centre, doi:10.5285/fb1e474d-456b-42a9-9a10-a02c35af10d2, 2013.

- Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., and Shaw, M. W.: Woodlands Survey Flora Data 1971–2001, NERC Environmental Information Data Centre, doi:10.5285/2d023ce9-6dbe-4b4f-a0cd-34768e1455ae, 2013.

All of the data sets are available from the CEH Environmental Information Data Centre Gateway (https://gateway.ceh.ac.uk) and via the following links: doi:10.5285/4d93f9ac-68e3-49cf-8a41-4d02a7ead81a, doi:10.5285/d6409d40-58fe-4fa7-b7c8-71a105b965b4, doi:10.5285/fb1e474d-456b-42a9-9a10-a02c35af10d2, doi:10.5285/2d023ce9-6dbe-4b4f-a0cd-34768e1455ae (see Table 5).

Data sets are provided under the terms of the Open Government Licence (http://eidchub.ceh.ac.uk/administration-folder/tools/ceh-standard-licence-texts/ceh-open-government-licence/plain, http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/).

The metadata are stored in the ISO 19115 (2003) schema (International Organization for Standardization, 2015) in the

UK Gemini 2.1 profile (UK GEMINI, 2015). Users of the data sets will find the following documents useful: "Long-term ecological change in British woodland (1971–2001)" (Kirby et al., 2005), "Woodlands Survey of Great Britain 1971–2001: dataset documentation" (Smart et al., 2013) (both supplied as supporting information with the data sets), "The effect of landscape-scale environmental drivers on the vegetation composition of British woodlands" (Corney et al., 2004) and the site reports, written by the Site Surveyors (2003).

7 Conclusions

The countryside of Great Britain and its woods have changed considerably over the last 50 years, for a variety of reasons. Some change has been gradual and can be attributed to factors such as evolving farming and forestry practices, climate change and atmospheric pollution. These have driven gradual responses in the composition and structure of woods. Other woods have undergone sudden change, in response to drivers such as the Dutch elm disease outbreak of the late 1960s and 1970s or the 1987 storm in south-east England.

The Woodland Survey of Great Britain thus provides a rare opportunity to explore the effects of a range of potential drivers of woodland change that operated between 1971 and 2001. It is a unique data set, consisting of a detailed range of ecological measurements at a national scale, covering a time span of over 30 years. It is also notable for the range of sites that have been revisited after such a long interval.

Author contributions. C. M. Wood prepared the manuscript with significant contributions from both co-authors, and is the current database manager for the Land Use Research Group at CEH Lancaster. S. M. Smart managed the survey in 2001 and has since carried out a wide range of analyses using the data sets described. R. G. H. Bunce designed the experiment (along with M. W. Shaw), ran the project in 1971 and made substantial contributions to the 2001 survey.

Acknowledgements. We thank H. I. J. Black, who managed the woodland soil analysis in 2001, assisted by G. Havelock, and also K. Kirby, who was the nominated contract officer for the re-survey and oversaw the production of the 2005 report. We also thank the land owners and agents who gave permission to conduct surveys on their land. Without their cooperation this project would not have been possible. Particular thanks also go to the funders: (1971) the Nature Conservancy; (2001) the Countryside Council for Wales, the Department of Environment, Food and Rural Affairs, English Nature, the Forestry Commission, the Joint Nature Conservation Committee, Scottish Natural Heritage, and the Woodland Trust; to the 2001 steering group: S. Bailley, M. Crick, J. Goodwin, K. Holl, D. Jackson, G. Patterson, S. Pryor, M. Smith, R. Smithers, A. Stott, and I. Strachan; to R. Isted and A. Wagstaff, who ran the contract for the 2001 resurvey; to the surveyors and support staff of 1971: R. Bunce, C. Barr, W. Bowen, A. H. F. Brown, A. D. Horrill, J. M. Sykes, N. E. Barber, P. A. Bassett, S. Daggitt, J. C. Holmes, I. P. Howes, T. J. Moss, G. K. A. Reynolds, P. L. Rye, C. Sargent, C. Smith, D. J. Taylor, and P. Wilkins; and to the surveyors and support staff of 2001: R. & F. Bunce, R. Cowan, L. Edey, H. Ford, T. Goucher, R. Hadden, V. Howard, M. King, K. Krowther, B. Lascelles, M. Prosser, J. Smith, H. Wallace, B. Wheeler, P. Wilson, S. M. Wilson, A. Scott, C. Hallam, D. Abel, G. Havelock, L. Maskell, M. Rossall, N. Ray, D. Singleton, and C. Wood. We thank the two anonymous referees, whose comments improved the content of this publication.

Edited by: F. Huettmann

References

Allen, S. E.: Chemical Analysis of Ecological Materials, 2nd Edn., Blackwell Scientific Publications, 1989.

Asner, G. P.: Organismic remote sensing for tropical forest ecology and conservation, Ann. Mo. Bot. Gard., 100, 127–140, doi:10.3417/2012016, 2015.

Avery, B. W.: Soil classification for England and Wales [higher categories], Technical Monograph, Soil Survey of England and Wales, 1980.

Balmford, A., Crane, P., Dobson, A., Green, R. E., and Mace, G. M.: The 2010 challenge: data availability, information needs and extraterrestrial insights, Philos. T. R. Soc. B, 360, 221–228, doi:10.1098/rstb.2004.1599, 2005.

Böhl, J. and Brändli, U.-B.: Deadwood volume assessment in the third Swiss National Forest Inventory: Methods and first results, Eur. J. For. Res., 126, 449–457, doi:10.1007/s10342-007-0169-3, 2007.

Bunce, R. G. H.: British woodlands in an European context, in: Forest and woodland ecology: an account of research being done in ITE, edited by: Last, F. T. and Gardiner, A. S., 7–11, (ITE symposium no. 8), Institute of Terrestrial Ecology, Cambridge, 1981.

Bunce, R. G. H.: A field key for classifying British woodland vegetation. Part 1, Institute of Terrestrial Ecology, Cambridge, 1982.

Bunce, R. G. H.: A field key for classifying British woodland vegetation. Part 2, London HMSO, 1989.

Bunce, R. G. H. and Jeffers, J. N. R.: Native pinewoods of Scotland, Institute of Terrestrial Ecology, Cambridge, 1977.

Bunce, R. G. H. and Shaw, M. W.: Classifying woodlands for nature conservation, Forestry and British Timber, 1, 23–25, 1972.

Bunce, R. G. H. and Shaw, M. W.: National survey of semi-natural woodlands, in: Merlewood Research Station Report for 1970–72, The Nature Conservancy, 1973a.

Bunce, R. G. H. and Shaw, M. W.: A standardised method for ecological survey, J. Environ. Manage., 1, 239–258, 1973b.

Bunce, R. G. H. and Smith, R. S.: An ecological survey of Cumbria, Kendal Cumbria County Council and Lake District Special Planning Board, 1978.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE land classification of Great Britain 1990, NERC Environmental Information Data Centre, doi:10.5285/ab320e08-faf5-48e1-9ec9-77a213d2907f, 1990.

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., and Howard, D. C.: The ITE Land Classification: providing an environmental stratification of Great Britain, Environ. Monit. Assess., 39, 39–46, doi:10.1007/978-94-009-1653-1_5, 1996.

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., Howard, D. C., Scott, W. A., Smart, S. M., Van de Poll, H. M., and Watkins, J. W.: Vegetation of the British countryside – the Countryside Vegetation System, Grange-over-Sands Institute of Terrestrial Ecology, ECOFACT Vol. 1, 1999.

Bunce, R. G. H., Metzger, M. J., Jongman, R. H. G., Brandt, J., de Blust, G., Elena-Rossello, R., Groom, G. B., Halada, L., Hofer, G., Howard, D. C., Kovář, P., Mücher, C. A., Padoa-Schioppa, E., Paelinx, D., Palo, A., Perez-Soba, M., Ramos, I. L., Roche, P., Skånes, H., and Wrbka, T.: A standardized procedure for surveillance and monitoring European habitats and provision of spatial data, Landsc. Ecol., 23, 11–25, doi:10.1007/s10980-007-9173-8, 2008.

Bunce, R. G. H., Bogers, M., and Roche, P.: Manual for habitat and vegetation surveillance and monitoring: temperate, mediterranean and desert biomes, Alterra Report No. 2154, Wageningen, 2011.

Carey, P. D., Wallis, S., Chamberlain, P. M., Cooper, A., Emmett, B. A., Maskell, L. C., McCann, T., Murphy, J., Norton, L. R., Reynolds, B., Scott, W. A., Simpson, I. C., Smart, S. M., and Ullyett, J. M.: Countryside Survey: UK Results from 2007, NERC/Centre for Ecology & Hydrology, Lancaster, 2008.

Cooke, R. and Kirby, K.: The use of a new woodland classification in surveys for nature conservation purposes in England and Wales, Arboricultural Journal, 18, 167–186, doi:10.1080/03071375.1994.9747013, 1994.

Corney, P. M., Le Duc, M. G., Smart, S. M., Kirby, K. J., Bunce, R. G. H., and Marrs, R. H.: The effect of landscape-scale environmental drivers on the vegetation composition of British woodlands, Biol. Conserv., 120, 491–505, doi:10.1016/j.biocon.2004.03.022, 2004.

Corney, P. M., Le Duc, M. G., Smart, S. M., Kirkby, K. J., Bunce, R. G. H., and Marrs, R. H.: Relationships between the species composition of forest field-layer vegetation and environmental drivers, assessed using a national scale survey, J. Ecol., 94, 383–401, doi:10.1111/j.1365-2745.2006.01094.x, 2006.

DeVries, B., Verbesselt, J., Kooistra, L., and Herold, M.: Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series, Remote Sens. Environ., 161, 107–121, doi:10.1016/j.rse.2015.02.012, 2015.

Emmett, B. E. and GMEP team: Glastir Monitoring & Evaluation Programme. First Year Annual Report to Welsh Government (Contract reference: C147/2010/11), NERC/Centre for Ecology & Hydrology, 442 pp., 2014.

Forestry Commission: Census of Woodlands 1947–1949. Woodlands of 5 acres and over, London, 264 pp., 1952.

Forestry Commission: Census of Woodlands 1965–70, Forestry Commission, London, 1970.

Forestry Commission: Census of Woodlands and Trees 1979–82, Forestry Commission, Edinburgh, 1984.

Forestry Commission: National Inventory of Woodland and Trees, Forestry Commission, Edinburgh, 2003.

Forestry Commission: National Forest Inventory, available at: http://www.forestry.gov.uk/website/forestry.nsf/byunique/infd-89s9ls, last access: 15 January 2013.

Groom, G. and Reed, T.: Strategic landscape monitoring for the Nordic countries, Nordic Council of Ministers, Copenhagen, 2001.

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

Henrys, P. A., Butler, A., Jarvis, S., Smart, S. M., and Fang, Z.: MultiMOVE Model: Ensemble niche modelling of British vegetation v2.0.1., NERC Environmental Information Data Centre, doi:10.5285/94ae1a5a-2a28-4315-8d4b-35ae964fc3b9, 2015.

Hill, M. O., Bunce, R. G. H., and Shaw, M. W.: Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland, J. Ecol., 63, 597–613, doi:10.2307/2258738, 1975.

Hinrichsen, D.: The forest decline enigma, BioScience, 37, 542–546, 1987.

International Organization for Standardization: ISO 19115:2003, available at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020, last access: 27 July 2015.

Kändler, G.: The design of the second German national forest inventory, in: Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium, Monterey, CA, 16–19 October 2006, edited by: McRoberts, R. E., Reams, G. A., Van Deusen, P. C., and McWilliams, W. H., Department of Agriculture, Forest Service, 19–24, General Technical Report WO-79, 404 pp., 2009.

Kandler, O. and Innes, J. L.: Air pollution and forest decline in Central Europe, Environ. Pollut., 90, 171–180, doi:10.1016/0269-7491(95)00006-D, 1995.

Keller, M. (Ed): Swiss National Forest Inventory. Manual of the field survey 2004–2007, Swiss Federal Research Institute WSL, Birmensdorf, 269 pp., 2011.

Kirby, K., Bines, T., Burn, A., Mackintosh, J., Pitkin, P., and Smith, I.: Seasonal and observer differences in vascular plant records from British woodlands, J. Ecol., 74, 123–131, doi:10.2307/2260353, 1986.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., and Smithers, R. J.: Long term ecological change in British woodland (1971–2001). A re-survey and analysis of change based on the 103 sites in the Nature Conservancy "Bunce 1971" woodland survey, Final report, Peterborough: English Nature (English Nature Research Reports Number 653), 139 pp. + Appendices, 2005.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., and Shaw, M. W.: Woodlands survey flora data 1971–2001, NERC Environmental Information Data Centre, doi:10.5285/2d023ce9-6dbe-4b4f-a0cd-34768e1455ae, 2013a.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., and Shaw, M. W.: Woodlands survey tree diameter data 1971–2001, NERC Environmental Information Data Centre, doi:10.5285/4d93f9ac-68e3-49cf-8a41-4d02a7ead81a, 2013b.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., and Shaw, M. W.: Woodlands survey soil data 1971–2001, NERC Environmental Information Data Centre, doi:10.5285/fb1e474d-456b-42a9-9a10-a02c35af10d2, 2013c.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., Smithers, R. J., and Shaw, M. W.: Woodlands survey site information 1971–2001, NERC Environmental Information Data Centre, doi:10.5285/d6409d40-58fe-4fa7-b7c8-71a105b965b4, 2013d.

Mandallaz, D.: Sampling techniques for forest inventories, CRC Press, Boca Raton, 2007.

Marrs, R., Kirby, K., Le Duc, M., McAllister, H., Smart, S., Oksanen, J., Bunce, R., and Corney, P.: Native dominants in British woodland-a potential cause of reduced species-richness?, New Journal of Botany, 3, 156–168, doi:10.1179/2042349713Y.0000000028, 2013.

Milner, C.: Shetland project monitoring report Nature Conservancy Council, Contract Report to the Nature Conservancy Council, Institute of Terrestrial Ecology, Grange-over-Sands, 1975.

National Forest Inventory: The Swiss National Forest Inventory, available at: http://www.lfi.ch/lfi/lfi-en.php, last access: 10 July 2015.

Ratcliffe, D. A.: A Nature Conservation Review: The Selection of Biological Sites of National Importance to Nature Conservation in Britain, Cambridge University Press, 1977.

Roche, P. and Geijzendorffer, I.: EBONE: integrated figures of habitat and biodiversity indicators: quantifying indicators of an integrated biodiversity observation system, Alterra, Wageningen UR, 2013.

Rodwell, J. S.: British Plant Communities: Volume 1, Woodlands and Scrub, Cambridge University Press, 1991.

Shaw, M. W. and Bunce, R. G. H.: National Woodlands Classification 1971 Handbook of Field Methods, Merlewood Research Station, Institute of Terrestrial Ecology, Grange-over-Sands, 1971.

Site Surveyors: Long term ecological change in British woodland (1971–2001). Site reports from the 103 sites surveyed in the Nature Conservancy "Bunce 1971" woodland survey and resurveyed in 2001–3, edited by: Rossall, M., Centre for Ecology and Hydrology, Lancaster, 2003.

Smart, S. M., Bunce, R. G. H., Black, H. I. J., Ray, N., Bunce, F., Kirby, K., Watson, R., and Singleton, D.: Measuring long term change in biodiversity in British woodlands (1971–2000) – a pilot re-survey of 14 sites from the ITE/NCC "Bunce" woodland survey and two from native pinewoods. Final report, English Nature, 56 pp. + Appendices, 2001.

Smart, S. M., Kirby, K., Corney, P., and Wood, C. M.: Woodlands Survey of Great Britain 1971–2001: dataset documentation, 2013.

Smart, S. M., Ellison, A. M., Bunce, R. G., Marrs, R. H., Kirby, K. J., Kimberley, A., Scott, A. W., and Foster, D. R.: Quantifying the impact of an extreme climate event on species diversity in fragmented temperate forests: the effect of the October 1987 storm on British broadleaved woodlands, J. Ecol., 102, 1273–1287, doi:10.1111/1365-2745.12291, 2014.

Smith, W. B.: Forest inventory and analysis: a national inventory and monitoring program, Environ. Pollut., 116, Supplement 1, S233–S242, doi:10.1016/S0269-7491(01)00255-X, 2002.

Sousa, A. M. O., Goncalves, A. C., Mesquita, P., and Marques da Silva, J. R.: Biomass estimation with high resolution satellite images: A case study of Quercus rotundifolia, ISPRS J. Photogramm., 101, 69–79, doi:10.1016/j.isprsjprs.2014.12.004, 2015.

Steele, R. C.: National survey of woodlands, reprinted from Society for the Promotion of Nature Reserves Handbook, Lincoln, 1968.

Sykes, J., Horrill, A., and Mountford, M.: Use of visual cover assessments as quantitative estimators of some British woodland taxa, J. Ecol., 71, 437–450, 1983.

Sykes, J. M. and Horrill, A. D.: Survey methods within woodland, Nature Conservancy Council, Peterborough (Contract HF3/03/107), 1979.

Thünen Institute: German National Forest Inventory, available at: http://www.ti.bund.de/en/wo/projects/monitoring-of-forests/bundeswaldinventur/national-forest-inventory/, last access: 10 July 2015.

UK GEMINI: UK GEMINI, available at: http://www.agi.org.uk/join-us/agi-groups/standards-committee/uk-gemini, last access: 12 May 2015.

United States Forest Service: Forest Inventory and Analysis National Program, available at: http://www.fia.fs.fed.us/, last access: 14 May 2015a.

United States Forest Service: National Forest Health Monitoring Program, available at: http://fhm.fs.fed.us/fact/index.htm, last access: 15 May 2015b.

Wani, A. A., Joshi, P. K., and Singh, O.: Estimating biomass and carbon mitigation of temperate coniferous forests using spectral modeling and field inventory data, Ecol. Inform., 25, 63–70, doi:10.1016/j.ecoinf.2014.12.003, 2015.

Williams, W. and Lambert, J. M.: Multivariate methods in plant ecology: I. Association-analysis in plant communities, J. Ecol., 47, 83–101, doi:10.2307/2257249, 1959.

Earth Syst. Sci. Data, 8, 177–189, 2016
www.earth-syst-sci-data.net/8/177/2016/
doi:10.5194/essd-8-177-2016

Open Access Earth System Science Data

Ecological survey of the native pinewoods of Scotland 1971

Claire M. Wood[1] **and Robert G. H. Bunce**[2]

[1]Centre for Ecology and Hydrology, Lancaster Environment Centre, Bailrigg, Lancaster LA1 4AP, UK
[2]Estonian University of Life Sciences, Kreuzwaldi 5, 51014 Tartu, Estonia

Correspondence to: Claire M. Wood (clamw@ceh.ac.uk)

Received: 3 February 2016 – Published in Earth Syst. Sci. Data Discuss.: 10 February 2016
Revised: 8 April 2016 – Accepted: 20 April 2016 – Published: 10 May 2016

Abstract. In 1971, a comprehensive ecological survey of the native pinewoods of Scotland was carried out by the Institute of Terrestrial Ecology. The survey was initiated as a consequence of growing concern about the status of the pinewood resource. Since the twentieth century, this unique habitat is widely recognised, not only by ecologists for its inherent biodiversity but also by the general public for its cultural and amenity value. The survey, utilising demonstrably repeatable methods, collected information on ground flora, soils, forest structure and also general site information from the major 27 sites of the 35 sites identified as truly native pinewoods in Scotland. The results from the survey prompted the organisation of an international symposium in 1975, which set the conservation agenda for the old Caledonian pinewoods. The data collected during the 1971 survey are now publicly available via the following DOI: doi:10/7xb ("Habitat, vegetation, tree and soil data from Native Pinewoods in Scotland, 1971"). Although the data are now 44 years old, the repeatable methods will allow for a resurvey to take place, in order to assess changes in the vegetation, habitats and tree composition in a statistically robust manner.

1 Introduction

Scots pine (*Pinus sylvestris* L.) is the most widely distributed conifer in the world (Steven and Carlisle, 1959). The only truly native and natural pinewoods in Britain are located in the Highlands of Scotland. Often referred to as the Caledonian Forest, the Scottish native pinewoods are of great interest, not only to the ecologist but also to the general public for cultural and amenity value. The native pinewoods of Scotland have a high biodiversity value and contain their own distinctive plants (and taxa). The pines themselves are genetically distinct, being of a unique variety (*Pinus sylvestris* var. *scotica*), which has adapted to the wetter and windier conditions of Scotland (Bain, 2013).

Concern for the ecological state of the native pinewoods was initiated in the 1940s and 1950s, resulting in a historical study undertaken by A. Carlisle in the late 1950s, and published as *The Native Pinewoods of Scotland* (Steven and Carlisle, 1959). This covered their location, condition and history, served to highlight the state of the native woodland resource and stimulated interest in the ecological value of the forests. Accordingly, a survey of the native pinewoods of Scotland was planned in 1970 to establish the range of variation within the pinewoods and to assess the state of the resource (Bunce, 1973). The survey was co-ordinated by the Institute of Terrestrial Ecology at Merlewood Research Station, Cumbria (now part of the Centre for Ecology and Hydrology), in conjunction with conservation colleagues in Scotland. The results of the survey were presented at an International Symposium held in Aviemore, Scotland, in 1975 (Bunce, 1977), and indicated a reduction in the area of the pinewoods since 1959. The survey and symposium served as an alert to stop taking the pinewoods for granted.

The survey was mainly carried out from 17 July to 24 August 1971 by graduates from the Forestry Department at Aberdeen University, after an introductory course held in Abernethy Forest (supervised by the second author). Thirty-five native pinewoods were identified in the book by Steven and Carlisle (1959), determined by criteria relating to historical records, age structure and the relative absence of human interference. Twenty-six of the 35 sites described by Steven and Carlisle were included in the 1971 survey (another site was surveyed in 1972 and the remaining eight mainly consisted of scattered trees). The methods followed those standardised by Bunce and Shaw (1973), described below, and also used in many subsequent surveys, including a national survey of semi-natural woodlands, also undertaken in 1971 (Wood et al., 2015b).

2 The native pinewood habitat

The terms "ancient" and "natural" tend to be used to describe woods originating before 1750, without subsequent human planting (in practice, woods in Britain are rarely unmanaged in some form or other, and therefore the term semi-natural is perhaps more appropriate) (Peterken, 1996; Balfour, 1977). Within the pinewoods, individual trees can live as long as 600 years, but generally live to around 250 (Bain, 2013).

As with other forests and woodlands in Britain, the pinewoods have been exploited by man over the centuries. The native Scottish pinewoods are the remaining fragments of an original forest system, reaching a maximum extent in Britain around 7500–4000 BP, as demonstrated by pollen records and macrofossil evidence (Bennett, 1984). Even sixteenth century maps provide evidence of a much wider distribution of Scots pine in the past than currently exists (Smout, 2006; Bain, 2013). Many of the remaining remnants have survived because they are remote or exist on poor soils. The remaining woods are therefore generally small, isolated and dominated by the one species of tree, the Scots pine.

Prior to the twentieth century, the pinewoods were depleted due to a range of human factors. Particularly during the seventeenth century, the pinewoods were progressively exploited as deer hunting forests, for livestock grazing, and as a timber and fuel resource (Smout, 2006; Steven and Carlisle, 1959; Gimingham, 1977). There is documented evidence that at least a dozen former areas of native pinewood have been lost since that time, although it is difficult to determine the exact details of all of the lost areas of pinewood due to incomplete historic records (Smout, 2006).

Some of the decline might also be attributed to historic climatic factors. As highlighted by Smout (2006), many lost pinewoods were located in the west of Scotland. As regeneration is known to occur more readily in the east, where it is much drier, it seems likely that wet and windy oceanic weather occurring in the sixteenth and seventeenth centuries would have had an adverse effect (Smout, 2006).

The consequences of these external influences on the pinewoods were threefold. Firstly, a reduction in area occurred, secondly a reduction in diversity occurred (*Betula* spp., *Sorbus aucuparia* and *Juniperus* spp. were all more abundant in the forests than now), and thirdly changes in the density and age structure of pure stands took place (Gimingham, 1977).

It was only during the twentieth century that the pinewoods started to be appreciated for their biodiversity value. Pinewoods are now recognised as an Annex I Priority Habitat under the EU Habitats Directive (JNCC, 2015; Romão, 2013) and are also listed on the Scottish Biodiversity List by the Scottish Government (Scottish Government, 2015). Nationally scarce plants may be found in the woods, such as twinflower (*Linnaea borealis*) and creeping lady's tresses (*Goodyera repens*). In addition to a unique ground flora, several varieties of native fauna are associated with the pinewoods, such as red squirrels (*Sciurus vulgaris*), the Scottish wildcat (*Felis silvestris*) and the pine marten (*Martes martes*). Bilberry (*Vaccinium myrtillus*) pollination depends on a rare bumblebee (*Bombus monticola*) (Bain, 2013; Lowe, 1977). The bird fauna is characteristic, and around 70 species of birds are known to breed regularly in the pinewoods. Several species are found in the pinewoods which are scarce elsewhere in Britain, including the crested tit (*Lophophanes cristatus*), the Scottish crossbill (*Loxia scotica*) and the Capercaillie (*Tetrao urogallus*) (Newton and Moss, 1977).

Broadleaved species are often an important component of the pinewoods, particularly birch (*Betula* sp.), oak (*Quercus* sp.), rowan (*Sorbus aucuparia*) and juniper (*Juniperus* sp.).

2.1 Survey sites

A set of 27 sites were chosen from 35 included as maps in the Steven and Carlisle book, as shown in Fig. 1 and Table 1. The chosen sites were the major areas of woodland, the remaining eight mainly consisting of scattered trees. The outlines of the forests were taken from this book, and 16 dispersed randomised points were marked and then located in the field by compass bearings and pacing from a recognised nearby landmark. Strict rules were imposed to ensure that there was no avoidable alteration to the predetermined position of the plots (Bunce, 1977).

2.2 Site descriptions

Steven and Carlisle (1959) divided the pinewoods into eight regional groups (as shown in Fig. 1). At least one site was surveyed from all of the groups mentioned.

The forests within the northern group are the isolated remnants of pinewood forests which would have grown throughout Sutherland. Overall, the pinewoods in this group are the smallest of the woods. Of these, only Amat was included in this survey. Records indicate that Amat would once have been part of extensive woodland that stretched across the valley, much of which was felled for shipbuilding in the eighteenth century (Steven and Carlisle, 1959).

The Deeside group, in Aberdeenshire, includes Glen Tanar, at the easternmost limit of the pinewoods in Scotland – as well as Ballochbuie and Mar. As far back as Queen Victoria (reigned 1837–1901), the royal family have taken an interest in the conservation of the pinewoods in this area; hence,

Table 1. List of surveyed pinewoods.

Site number	Name	OS grid ref.	Area of pure pine (ha) (Steven and Carlisle, 1959)
1	Glen Tanar	NO459941	863
2	Ballochbuie	NO200897	860
3	Mar	NO099892	415
4	Abernethy	NH970183	827
5	Rothiemurchus	NH925065	932
6	Glenmore	NH981091	65
7	Glen Feshie	NN841939	288
8	Rannoch	NN582542	372
9	Meggernie	NN553467	131
10	Glen Moriston	NH313117	80
11	Glengarry	NH226003	324
12	Barrisdale	NG883071	183
13	Loch Arkaig and Glen Mallie	NN085900	164
14	Ardgour	NM986750	131
15	Glen Affric	NH203225	710
16	Glen Cannich	NH227315	302
17	Glen Strathfarrar	NH238375	204
18	Guisachan and Cougie	NH318241	255
19	Coulin	NG997568	142
20	Achnashellach	NH028483	130
21	Shieldaig	NG825522	51
22	Amat	NH454895	98
23	Loch Maree	NG881727	160
24	Black Mount	NN284441	55
25	Glen Orchy	NN240330	55
26	Tyndrum	NN328279	66
27	Dulnain	NH831185	*Not given*

there has been relatively little felling of the forest since that time (Bain, 2013).

The Speyside group includes some of the most extensive remaining pinewoods: Abernethy and Rothiemurchus, as well as Glenmore, Glen Feshie and Dulnain. These woodlands are now within the Cairngorms National Park, and are much visited, particularly the former three. Thanks to regeneration and restoration in these woods, they demonstrate good examples of Caledonian forest. In the nineteenth century, Rothiemurchus and Abernethy were hunting forests (Smout, 2006; Bain, 2013).

The Rannoch group, lying in the heart of Perthshire, includes the Black Wood of Rannoch and the Old Wood of Meggernie. Birch is an important constituent of both of these woods.

The pinewoods in the Great Glen group are situated in the valleys of the glen, and include Glen Moriston, Glengarry, Barrisdale, Loch Arkaig and Glen Mallie, and Ardgour. The forests here form part of a landscape of open moors and mountains but have been much influenced by planting of exotic conifers.

In the Wester Ross group, at the western edge of the natural Scots pine entire world range, is Shieldaig, near the Atlantic coast beyond the isolated Torridon mountains. The group also includes Loch Maree, Coulin and Achnashellach. The region has a high rainfall (annual total in 2015: 2254 mm) relative to Britain as a whole (annual total in 2015: 1289 mm), and in comparison to the east of Scotland (annual total in 2015: 1368 mm) (Met Office, 2016). The surviving pinewoods are relatively small in area and generally surrounded by moorland.

The southern group consists of relatively small remnants east of Loch Linnhe and the Firth of Lorne including Black Mount Woods, Glen Orchy and Tyndrum.

The Strath Glass group, located in an area of the Highlands popular with visitors, includes Glen Affric, Glen Cannich, Glen Strathfarrar, and Guisachan and Cougie.

The areas of the different sites vary considerably, and measurements can only be approximate as there are difficulties in determining the exact areas. For example, it is difficult to determine the point where a forest becomes moorland with scattered trees and whether areas of bog within the woodlands should be included. This can partly account for the reason why different estimates of area are available from different sources, as described below. Areal estimates are provided in Goodier and Bunce (1977), as measured from the

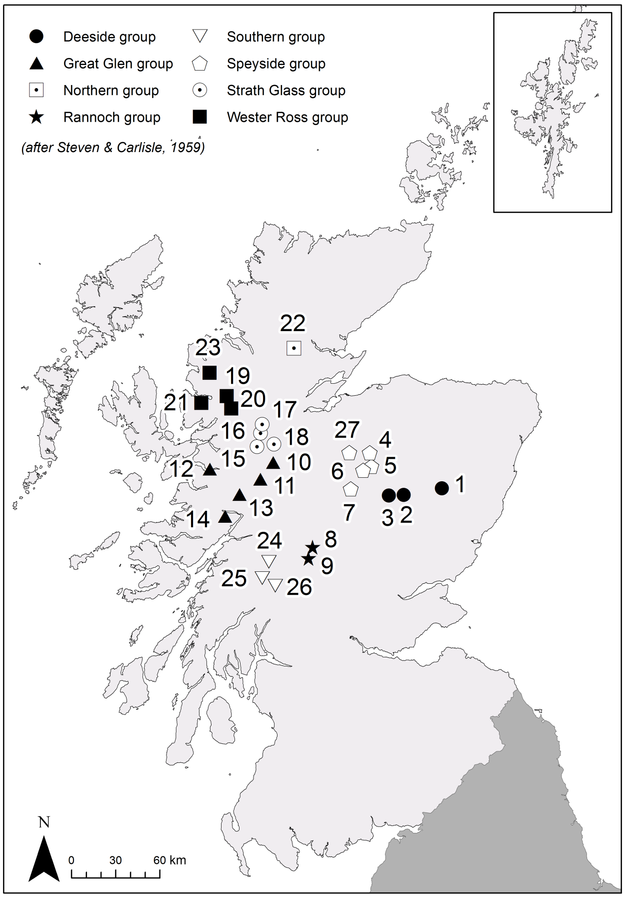


Figure 1. Map of survey site locations.

maps given in Steven and Carlisle (1959) and shown in Table 1. These figures could now be improved by interpretation of modern aerial photography, with forest being at least 30 % cover (otherwise they should be considered as scattered trees) or a resurvey could perhaps provide improved area estimates based on a standard basal area per hectare. The largest of the pinewood sites are Glen Tanar, Abernethy, Rothiemurchus and Ballochbuie, all estimated as having over 800 ha of pure pine area. The smallest, with less than 60 ha of pure pine, are Glen Orchy, Black Mount and Shieldaig. The rest of the sites are estimated as having a range from between 65 and 710 ha of pure pine (although these figures must be treated with caution). The overall area of native pinewood as surveyed in the Native Woodland Survey of Scotland (NWSS) is 87 599 ha (Patterson et al., 2014). This figure is greatly in excess of the pinewoods described as historically certain to be of native origin by Steven and Carlisle (1959). Goodier and Bunce (1977) measured the woodland areas from the maps in Steven and Carlisle (1959) which gave an overall figure of 10 700 ha, although the area of relatively dense pine from the sample survey described in this paper was only 1600 ha. Whilst the larger figure will contain sites comparable to the true old Caledonian pinewoods, they do not have the necessary criteria to be considered as old-growth forests and also do not satisfy the definition given in Annex I of the Habitats Directive (Romão, 2013).

2.3 Plot layout and descriptions

Following the methodology of Bunce and Shaw (1973), 16 plots were randomly positioned within each site and the location of each was marked on a 1 : 25 000 map. Each plot was 14.1 × 14.1 m (200 m^2) (Fig. 2) and constructed as shown in Fig. 3, with one centre post and four corner posts, with a set of four strings tagged with markers at specified distances. The centre post had a right angled gauge affixed to the top, in order to orientate the plot at random. In the field, plots were located by pacing from the nearest relocatable feature. As described below, data were then collected on ground flora, tree and shrub layers and soils. Habitat characteristics were recorded for the both the individual plots and each wood as a whole.

The sampling intensity of 16 plots per site was used in a concurrent survey, the Woodland Survey of Great Britain in 1971 (Wood et al., 2015b), and was chosen on the basis of previous experiences in surveying a wide variety of sites in the north of England and Wales. It also coincided with the time and manpower available (Bunce and Shaw, 1973).

3 Data collected

A range of data were collected within each of the surveyed pinewoods, as shown in Table 2, including ground flora records, tree information, shrubs and sapling information, seedling information, plot description and habitats, soil data and a whole wood description.

3.1 Site Information, plot locations and information, slope and aspect

For both the whole woodland site and for each of the 16 200 m^2 plots within, the presence and absence of a series of attributes were recorded (as summarised in Table 2). Attributes included management factors such as the presence of dead trees or stumps; physical factors such as the presence of rock or cliffs; habitat-related factors such as the presence of rotting stumps or hollow trunks; aquatic habitats; presence of buildings or open habitats such as glades and rides; presence of epiphytes on trees, animals and birds; and boundary types and nearby land use. A full list of habitats may be found in the field handbook (Shaw and Bunce, 1971) (supplied as supporting documentation with the data sets). The slope of each plot was measured in degrees using a hypsometer and the aspect of each plot was measured using a magnetic compass.

Table 2. Summary of data collected.

Data category	1971 survey
Ground flora	Species present in the plot % cover/abundance estimates, bryophyte collection
Trees	DBH (diameter at breast height) and species recorded from all four quadrants of the plot
Shrubs and saplings	DBH and species recorded from diagonally opposite quarters of the plot
Seedlings	Included with the ground flora records
Plot description and habitats	Tick list of features (broad categories): Tree management Regeneration Dead trees Epiphytes Rock habitats Aquatic habitats Open habitats Human elements Vegetation structure Animal signs
Soil data	Tick list description from small pit and augur boring in the centre of the plot – to determine soil type Composite soil sample from top 10–15 cm.
Whole wood description	Tick list of features (broad categories). As for plot, plus adjacent land use and boundary type

3.2 Vegetation data

Within the plot described in Fig. 2, the area within the first nest of the plot (2 × 2 m) was searched for the presence of all vascular plants (monocotyledons, dicotyledons and ferns), including tree species. This procedure was repeated for each nest of the quadrat, increasing the size each time as shown in Fig. 2. In the final nest (the whole 200 m^2 plot), the percentage cover (to the nearest 5 %) of each species was estimated. In addition, the total cover of bryophytes was estimated from the entire plot, as was an overall estimate for litter, wood, rock, bare ground and standing water. Bryophytes and lichens were collected separately and specimens identified later.

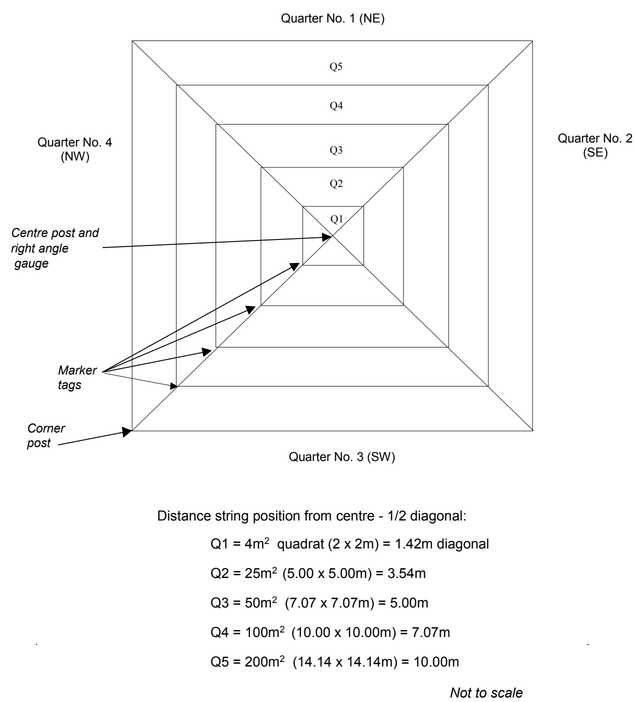


Figure 2. Plot layout.

3.3 Soil data

Soil samples were taken from every accessible plot in every woodland. A single composite soil sample was taken from each plot, at the centre of the vegetation quadrat, using a trowel. Samples (weighing approximately 1 kg) were taken to a depth of 15 cm and placed in a labelled plastic bag. On return to the laboratory, all soil samples were stored at 4 °C prior to processing and analyses. Soil samples were sieved using a 2 mm sieve. A pH reading was taken on a representative fresh subsample from each soil sample before air-drying

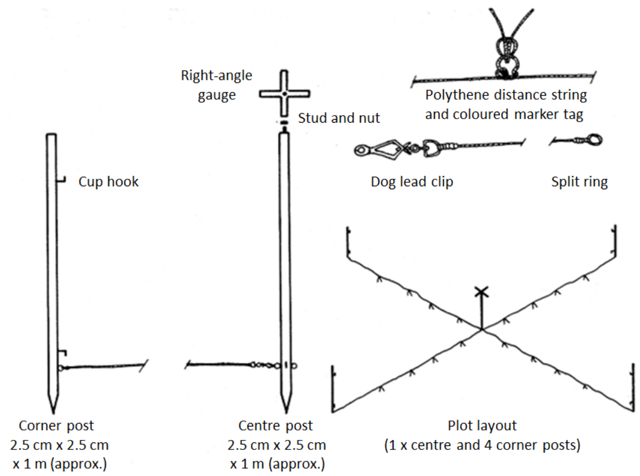


Figure 3. Plot construction.

at 20 °C. Unless otherwise stated, soil pH values in the data set are from the soil samples prior to air-drying ("fresh"). All analyses were carried out under the supervision of the Environmental Chemistry Section at the Institute of Terrestrial Ecology, Merlewood, following standard methodologies and quality control procedures (Allen, 1989), including the analyses of certified standard reference samples within batches.

3.4 Tree diameter

Trees, saplings and shrubs were recorded in the 200 m^2 plot, as described above. Decisions as to whether individuals are in the plot or not were based on the rooted base being 50 % or more within the plot.

For trees (stems of more than 5 cm diameter at breast height (DBH) of any species normally capable of attaining a treelike habit in Britain), the species and DBH of all stems in the whole plot greater than 5 cm were measured. Trees with multiple stems had each stem recorded separately. Standing dead trees were also measured and identified as such.

Saplings (definition as for trees, but with a height of less than 130 cm and with a DBH less than 5 cm) were recorded only in quarters 1 and 3 of the plot (see Fig. 2). The same measurements as for trees were made. Shrubs, like saplings, were also only recorded in quarters 1 and 3, and again the same measurements were taken. Shrubs were defined as species including hazel and juniper.

3.5 Data quality

During the survey, all survey teams were initially accompanied by a supervisor and regular visits into the field were made by the project leader to ensure consistency and quality in data recording according to criteria laid out in the field handbook (Shaw and Bunce, 1971).

The data sets were transferred from the original field sheets to spreadsheets in the 2000s. They were checked and corrected to produce a final validated copy. Standard validation checks included plot and site counts to ensure no duplicate numbering and hence double counting of plots; also, range checks were undertaken where possible for values falling within certain ranges, such as soil pH or slope values.

In terms of the soil data, descriptive profile data were collected to the standards set out in the training and field handbook, but they were not formally checked for quality aside from checks from supervisors during the survey. The soil pH was analysed using quality control measures as outlined in Allen (1989). These included the analyses of certified standard reference samples within batches.

4 Summary of findings: vegetation and general habitats

The ground flora vegetation gives a good indication of the state of the general environment, often more so than tree composition (Hill et al., 1975). In analysing the data, it is possible to focus on three levels: species level, plot level and site level. Overall, the dominant species found in the survey, as shown in Table 3, are revealed to be heather (*Calluna vulgaris*), bilberry (*Vaccinium myrtillus*), tormentil (*Potentilla erecta*), wavy hair grass (*Deschampsia flexuosa*) and cowberry (*Vaccinium vitis-idaea*). Creeping lady's tresses (*Goodyera repens*), a rare orchid only found in the Scottish pinewoods, was recorded 14 times.

The majority of the species recorded are not associated particularly with pinewoods. This indicates that the species composition, as expressed by the most frequent contributors, reflects the open nature of the forests and the frequency of other habitats such as bog surfaces (Bunce, 1977). The species composition often relates to upland heath vegetation, which is mostly derived from former extensive forests. As the pinewoods have retreated, the heathlands have extended in the drier parts of Scotland. In the west, there is a similar relationship with peat-forming vegetation (Gimingham, 1977).

The results can be compared to the results shown from the National Woodland Survey (Wood et al., 2015b; Kirby et al., 2005), as the same survey techniques were applied. Although virtually all the species recorded in the Native Pinewoods Survey were also found in the national survey, many were at low frequencies. However, many of the species in the national survey were absent from the pinewoods, reflecting the relatively limited range of variation within this habitat (Bunce, 1977).

Around 25 % of the pinewood sites also had exotic tree species planted. Many of these have now been felled because of a change in policy, but it will be many years before the ground vegetation recovers. A repeat survey would provide figures of the actual extent and impact of this felling.

Table 3. List of top 25 ground flora species recorded.

	Species	Common name	No. of records
1	*Calluna vulgaris*	Heather	398
2	*Vaccinium myrtillus*	Bilberry	356
3	*Potentilla erecta*	Tormentil	333
4	*Deschampsia flexuosa*	Wavy hair grass	299
5	*Vaccinium vitis-idaea*	Cowberry	287
6	*Molinia caerulea*	Purple moor-grass	286
7	*Blechnum spicant*	Hard fern	272
8	*Betula* sp.	Birch	233
9	*Agrostis canina*	Brown bent	214
10	*Pteridium aquilinum*	Bracken	212
11	*Erica tetralix*	Cross-leaved heather	210
12	*Carex echinata*	Star sedge	195
13	*Sorbus aucuparia*	Rowan	194
14	*Galium saxatile*	Heath bedstraw	186
15	*Narthecium ossifragum*	Bog asphodel	169
16	*Luzula multiflora*	Heath woodrush	156
17	*Erica cinerea*	Bell heather	145
18	*Pinus sylvestris*	Scots pine	142
19	*Viola riviniana/reichenbachiana*	Common dog violet	137
20	*Melampyrum pratense*	Common cow-wheat	130
21	*Polygala serpyllifolia*	Heath milkwort	127
22	*Succisa pratensis*	Devil's bit scabious	127
23	*Carex panicea*	Carnation sedge	115
24	*Carex binervis*	Green-ribbed sedge	113
25	*Eriophorum vaginatum*	Hare's-tail cottongrass	111

On a plot level, the design of the survey methods allows the vegetation to be classified into relatively homogenous groups. Using indicator species analysis (Hill et al., 1975), a key was able to be prepared, differentiating the major plot types. The full key may be viewed in Bunce (1977). In total, eight distinct plot types were differentiated (as summarised in Table 4). Each of the types shows variation in soil type and pH, slope and habitat types. In terms of comparison with previously recognised associations, plot type one is the least heterogeneous and can be recognised as approximately corresponding with the *Pinetum Hylocomieto–Vaccinetum* community, identified by McVean and Ratcliffe (1962). This was described as characteristic of moderately dense pinewood throughout the Central and Northern Highlands. Tall shrubs are generally absent, and *Goodyera repens* is exclusive to this association. The total number of species is not high, with characteristic species including *Pinus sylvestris*, *Calluna vulgaris*, *Vaccinium myrtillus*, *Vaccinium vitis-idaea* and *Hylocomium splendens*.

The other seven identified plot types are harder to compare. McVean and Ratcliffe (1962) identified a second community, *Pinetum Vaccineto–Callunetum*, characteristic of the more open forests, often with pine–birch mixtures and even pure birch woodland where this had colonised former pine ground. It differs from the first by being dominated by tall heather and *Vaccinium myrtillus* with deep sphagnum tussocks. This association encompasses a range of plot types, as identified in this 1971 survey.

A trend can be identified from the data in terms of environmental correlations with the different plot types. The highest correlations were found to be with peat depth. Other significant correlations were with the depth of parent material, depth of podzolic horizon, slope and the depth of the mixed/mineral horizon. Other correlations are with soil pH and soil type.

As many practical conservation problems are found at a site scale, it is important to analyse the data at a site level. In analysing the site data overall, four distinct site types were identified, which correlate strongly with their geographical distribution. These were classified as eastern, central, north western and south-western. Each type showed distinct ground flora composition, pine frequency, and age of trees (in terms of shrubs/sapling/trees) (Bunce, 1977).

Summary of findings: forest structure

Having statistically robust data providing information on the forest structure is useful in explaining the status of the resource. Few data sets exist regarding this, covering all of the native Scottish pinewoods (Mason et al., 2007). The Scottish pinewoods are characterised by a diverse structure, with irregular tree spacings, shapes and sizes. These variations oc-

Table 4. Summary of plot types identified by indicator species analysis.

Plot type	Description	No. of plots	% of plots with no trees	Mean altitude (m)	Mean slope (degrees)	Mean soil pH
Type 1	*Empetrum nigrum/Calluna vulgaris*	60	12	296	13	3.9
Type 2	*Vaccinium vitis-idaea/Calluna vulgaris*	66	21	269	14	4.1
Type 3	*Oxalis acetosella/Pteridium aquilinum*	60	21	271	12	4.5
Type 4	*Cirsium vulgare/Pteridium aquilinum*	42	46	233	21	4.9
Type 5	*Drosera rotundifolia/Molinia caerulea*	63	69	213	10	4.4
Type 6	*Erica tetralix/Calluna vulgaris*	57	21	254	10	4.0
Type 7	*Narthecium ossifragum/Molinia caerulea*	37	46	266	8	5.1
Type 8	*Solidago virgaurea/Molinia caerulea*	31	46	215	12	4.6

cur largely as a consequence of the range of slopes and soil types on which the pines grow, ranging from freely drained thin podzol soils on steep rocky crags to wet peaty ground (Bain, 2013). Data from this survey provide information on the structure of the tree layer, including tree diameters at breast height and tree densities. Initial analysis of the forest structure data was undertaken by Goodier and Bunce (1977) and is summarised as follows.

The diameter distributions for trees in individual sites (Fig. 4) show that the forests consist of largely older trees with few younger specimens, as Steven and Carlisle (1959) indicated. Some forests, in particular Loch Maree and Abernethy have reasonable diameter distributions, whereas others, such as Ballochbuie and Mar, are skewed towards the older larger classes. Glen Moriston and Glengarry require particular comment. Much of the forest at these two sites has been planted with *Picea sitchensis* and the younger trees are regenerating in small gaps among the plantations, suggesting the effects of the removal of grazing.

The structure of the forests does not depend solely upon the diameter distribution of the trees but also on their height and on the other species present. The forests in the east tend to have taller and probably more vigorous trees, whereas those in the west contain more stunted trees because of the factors associated with the depth of peat on the sites. The proportion of birch also varies between sites, and this has a critical effect on the populations of birds and insects (Goodier and Bunce, 1977).

5 The survey in context

The data from this survey are unique, as they provide a comprehensive, repeatable set of data giving information on many aspects of the pinewood ecosystem, and can be analysed in a statistically robust way. Information from surveys pertaining to the Scottish pinewoods predating this one tends to have been collected in a unrepeatable fashion or has not been as fully comprehensive in terms of the range of data collected (Smith, 1900; McVean and Ratcliffe, 1962; Steven and Carlisle, 1959). Subsequent surveys relating to the pinewoods have been on a smaller, localised scale, focused on specific sites (Mchaffie et al., 2002; Wilson and Puri, 2001; Vickers and Palmer, 2000) and are often focused on one particular aspect of the ecosystem, such as a particular species (often not restricted to the pinewood habitat) (Wilkinson et al., 2002; Summers and Buckland, 2011), forest structure (Mason et al., 2007; Summers et al., 1997) or regeneration (Scott et al., 2000; Baines et al., 1994; Palmer and Truscott, 2003). Some studies solely focus on areal extent (Roberts et al., 1992; Cameron et al., 2000). Bain (1987) carried out a geographically comprehensive review of the extent and condition of the native pinewoods, but this did not include information from vegetation or soil plots.

A major survey was carried out from 2006 to 2013 by the Forestry Commission Scotland, the Native Woodland Survey of Scotland (NWSS) (Patterson et al., 2014). Whilst this provides a comprehensive view of the extent, structure and condition of the tree species, the survey did not include comprehensive ground flora or soil assessments, limiting the assessment of the entire ecosystem. A repeat of the 1971 survey would enable assessment of changes within the woodlands to be made.

The results from this 1971 survey were presented at a symposium held in Aviemore, Scotland, in 1975, organised by the Institute of Terrestrial Ecology (Bunce and Jeffers, 1977). The conference raised awareness of the ecological value of the resource, and set the conservation agenda. By the end of the 1970s, 21 of the 35 Steven and Carlisle sites had been designated as sites of special scientific interest (Bain, 2013). In 1977, the Forestry Commission introduced a native pinewood grant scheme, whereby landowners were given grants to restore native pinewoods. By the 1990s, over 80 % of the native pinewoods were within protected sites, putting an end to major losses from felling and non-native planting (Bain, 2013).

6 Methodology in context

Together with the National Survey of Semi-Natural Woodlands (Wood et al., 2015b), this survey was the first time

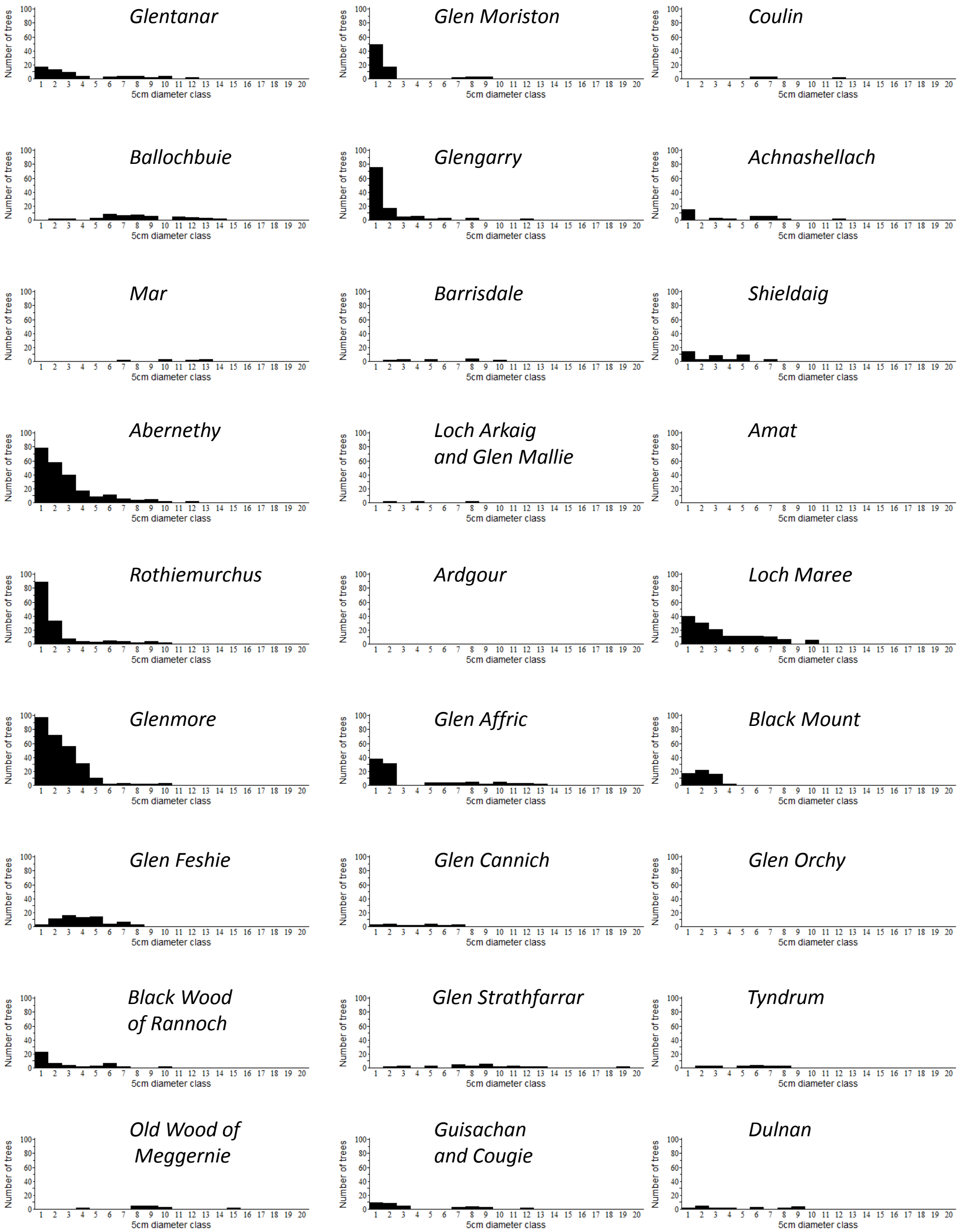


Figure 4. Diameter distribution of trees within each pinewood site.

that stratified random samples were being used to obtain an integrated assessment of the response of vegetation to the environment across a defined population. The structure of the project provided the basis for the further development of strategic survey methods. The methods used in the survey, originally described in Bunce and Shaw (1973), were intended to be comparable to the National Survey of Woodlands, also taking place in 1971 (Kirby et al., 2005; Wood et al., 2015b). Whereas the national survey had to be confined to a sample of British woodlands, the Scottish Pinewoods Survey aimed to be as exhaustive as possible. The methods utilised within these surveys were repeated successfully in subsequent regional surveys during the 1970s such as the Cumbria Survey (Bunce and Smith, 1978) and the Terrestrial Survey of Shetland (Milner, 1975).

Variations of the method using the concept of a woodland site, and subsequently a 1 km square sampled at random, with random plots sampled within, have become a standard sampling strategy used very successfully in several other large ecological surveys in Britain, such as the Countryside Survey (Carey et al., 2008), and the Glastir Monitoring and Evaluation Programme (Emmett and GMEP team, 2014). Outwith Great Britain, methods adapted from the basic principles in this survey have been developed to roll out across the whole of Europe as part of the European Biodiversity Observation Network (EBONE) project (Bunce et al., 2008, 2011). The methods were widely tested across 12 European countries, and also Israel, Australia and South Africa. The methods were proven to be robust, reliable and repeatable at a continental, landscape scale (Roche and Geijzendorffer, 2013).

A key aim of the sampling design was that the methods chosen should be standardised, and therefore repeatable. For the purposes of sampling in woodland, a large quadrat is necessary, both in order to include a reasonable number of trees and, if the canopy is dense, to accommodate an adequate sample of the ground vegetation. The size of the plot was chosen with reference to continental phytosociologists who at the time most widely used plots of between 100 and 200 m^2 (Bunce and Shaw, 1973). After preliminary field tests, it was found that the number of species recorded usually stabilised at this size. The area of 200 m^2 was thus adopted for this survey. As the focus of the survey is on ground flora as well as tree and shrub information, the square plot with inner nests aids a systematic search of the vegetation within the plot. It is also straightforward to layout in the field, and ensures a standard sized plot is laid out every time. For these reasons, the square plot was considered more advantageous than a circular plot. Plotless sampling was also dismissed, as it is not a suitable method for recording ground vegetation, only tree density. Random sampling was preferred to systematic sampling in this case to avoid the possibility of resonance with environmental features, for example a map grid line following the course of a stream. Dispersed random sampling also has practical advantages over systematic sampling, which requires continuous scale adjustment in order to obtain a constant sample from variable-sized areas (Bunce and Shaw, 1973). For the purposes of devising conservation policies, larger landscape units must be considered. Accordingly, data from the 16 random plots may be used to assess the vegetation for whole sites.

In terms of the repeatability of the survey, statistical analyses of temporal vegetation change are clearly more powerful when based on records from plots located in the same place rather than randomised to new locations for each survey. Surveys using the Bunce and Shaw (1973) methodology have been proven to be effectively repeatable. The British Countryside Survey of 1990 aimed to repeat vegetation plots first recorded in 1978, using the same information as would be available for the pinewood plots (plot maps and descriptions). The locational accuracy of the plot locations in 1990 is assessed in Prosser and Wallace (1992) and Barr et al. (1993). Overall, plots in this larger survey had a relocation rate of 87 %. Ideally, additional information such as plot photos and permanent plot markers would be introduced in a resurvey, as has been the case in the Countryside Survey to increase the repeatability.

Further analyses regarding plot repeatability were undertaken using data from the Woodland Survey of Great Britain, carried out in 1971 and again in 2001 and again using exactly the same Bunce and Shaw (1973) methodology as the Scottish Pinewoods Survey. In the repeat survey, the field surveyor relied only on the marked point on a map as the sole aid to relocating the 1971 plot location. It would be expected that, having made an effort to move near to the mapped point, the plot records from the repeat survey will, on average, be more similar to the respective 1971 plot record than if a completely new, random set of locations were chosen. Even if vegetation change occurs, species compositional data recorded from the same point at times 1 and 2 will tend to be more similar than data recorded from two random points at times 1 and 2. Whilst it is impossible to measure the amount of relocation error by exploiting a "true" set of temporal pairs (known to have been recorded in exactly the same position), it is possible to compare the average species compositional similarity between the ostensibly true temporal pairs with the average similarity for a random pairing of the 1971 data with the 2001 data. If, on average, attempts to relocate the true 1971 position had been successful then the similarity between the true pairs should be greater than the random pairs. Overall, at 94 % of the woodland sites, mean similarity was greater between "relocated" plot pairs compared to random-pair comparison, and for 57 % of sites the difference was significantly greater. A full account of this is given in Appendix 3 of Kirby et al. (2005).

7 Conclusions

The recently published data from the ecological survey of native Scottish pinewoods carried out in 1971 provide a comprehensive view of the pinewood habitat state in that year. Consequently, a detailed range of ecological measurements are now publicly available for all of the major native pinewoods identified by Steven and Carlisle (1959). The standardised methods allow for the possibility of a repeat of the survey, which would reveal changes in the condition, extent and composition of this unique habitat. This would provide the opportunity to explore a range of causal factors and drivers of change, such a grazing influences, management, climate change and pollution. It would be hoped that the range of mitigation factors introduced since the 1970s (such as the removal of exotic conifers) would have increased the extent, and improved the condition of this ecologically valuable resource.

Data availability

The data sets (Bunce et al., 2015) have been assigned digital object identifiers and users of the data must reference the data as follows:

Bunce, R. G. H., Shaw, M. W., and Wood, C. M.: Habitat, vegetation, tree and soil data from Native Pinewoods in Scotland, 1971, NERC Environmental Information Data Centre, doi:10.5285/56a48373-771c-4d4a-8b5a-45ef496c6e55, 2015. (Bunce et al., 2015)

This can be downloaded from the CEH Environmental Information Platform (https://eip.ceh.ac.uk/) from the following link: doi:10/7xb.

The data are provided under the terms of the Open Government Licence (http://eidchub.ceh.ac.uk/administration-folder/tools/ceh-standard-licence-texts/ceh-open-government-licence/plain, http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/).

The metadata are stored in the ISO 19115 (2003) schema (International Organization for Standardization, 2015) in the UK Gemini 2.1 profile (UK GEMINI, 2016).

Users of the data will find the following documents useful: Shaw and Bunce (1971) and Wood et al. (2015a) (both supplied as supporting documentation with the data sets).

Author contributions. Claire M. Wood prepared the manuscript with significant contributions from the co-author, and is the current database manager for the Land Use Research Group at CEH Lancaster. Robert G. H. Bunce designed the experiment (along with M. W. Shaw) and managed the project in 1971.

Acknowledgements. The fieldwork for the survey was carried out by K. Wilson, M. J. Bottomly, R. H. Reid, K. H. Chorlton, M. E. Ball, M. W. Shaw and D. J. Taylor. C. Helliwell, J. Brocklebank, and C. Barr provided help in the laboratory. M. W. Shaw provided help with developing computing methods, and C. J. Hallam assisted with data entry. Thanks are due to the many owners who gave permission for the site visits and to Forestry Commission and Nature Conservancy staff who also provided helpful support during the survey in 1971. We thank the two referees, L. Iverson and B. Young, for their constructive comments which have improved the quality of the manuscript.

Edited by: F. Huettmann

References

Allen, S. E.: Chemical Analysis of Ecological Materials, 2nd Edn., Blackwell Scientific Publications, Oxford, 1989.

Bain, C.: Native Pinewoods in Scotland: A Review 1957–1987, Royal Society for the Protection of Birds, Edinburgh, 1987.

Bain, C.: The Ancient Pinewoods of Scotland, Sandstone Press Ltd, Dingwall, 2013.

Baines, D., Sage, R., and Baines, M.: The implications of red deer grazing to ground vegetation and invertebrate communities of Scottish native pinewoods, J. Appl. Ecol., 31, 776–783, doi:10.2307/2404167, 1994.

Balfour, J.: Introduction to the symposium on Scottish native pinewoods, in: Native Pinewoods of Scotland, edited by: Bunce, R. G. H. and Jeffers, J. N. R., Institute of Terrestrial Ecology, Cambridge, 1977.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Fuller, R. M., Furse, M. T., Gillespie, M. K., Groom, G. B., Hallam, C. J., Hornung, M., Howard, D. C., and Ness, M. J.: Countryside Survey 1990: main report, Countryside 1990 Vol. 2, London: Department of the Environment, 1993.

Bennett, K.: The post-glacial history of Pinus sylvestris in the British Isles, Quaternary Sci. Rev., 3, 133–155, doi:10.1016/0277-3791(84)90016-7, 1984.

Bunce, R. G. H.: Native pinewood survey, in: Merlewood Research Station Report for 1970–72, The Nature Conservancy, Grange-over-Sands, 1973.

Bunce, R. G. H.: The range of variation within the pinewoods, in: Native pinewoods of Scotland, edited by: Bunce, R. G. H. and Jeffers, J. N. R., Institute of Terrestrial Ecology, Cambridge, 10–25, 1977.

Bunce, R. G. H. and Jeffers, J. N. R.: Native pinewoods of Scotland, Institute of Terrestrial Ecology, Cambridge, 1977.

Bunce, R. G. H. and Shaw, M. W.: A standardised method for ecological survey, J. Environ. Manage., 1, 239–258, 1973.

Bunce, R. G. H. and Smith, R. S.: An ecological survey of Cumbria, Structure Plan Working Paper, 4, Cumbria County Council and Lake District Special Planning Board, Kendal, 1978.

Bunce, R. G. H., Metzger, M. J., Jongman, R. H. G., Brandt, J., de Blust, G., Elena-Rossello, R., Groom, G. B., Halada, L., Hofer, G., Howard, D. C., Kovář, P., Mücher, C. A., Padoa-Schioppa, E., Paelinx, D., Palo, A., Perez-Soba, M., Ramos, I. L., Roche, P., Skånes, H., and Wrbka, T.: A standardized procedure for surveillance and monitoring European habitats and provision of spatial data, Landsc. Ecol., 23, 11–25, doi:10.1007/s10980-007-9173-8, 2008.

Bunce, R. G. H., Bogers, M., and Roche, P.: Manual for habitat and vegetation surveillance and monitoring: temperate, mediterranean and desert biomes, Alterra, Wageningen, 2011.

Bunce, R. G. H., Shaw, M. W., and Wood, C. M.: Habitat, vegetation, tree and soil data from Native Pinewoods in Scotland, 1971, NERC Environmental Information Data Centre, doi:10.5285/56a48373-771c-4d4a-8b5a-45ef496c6e55, 2015.

Cameron, A., Miller, D., Ramsay, F., Nikolaou, I., and Clarke, G.: Temporal measurement of the loss of native pinewood in Scotland through the analysis of orthorectified aerial photographs, J. Environ. Manage., 58, 33–43, doi:10.1006/jema.1999.0306, 2000.

Carey, P. D., Wallis, S., Chamberlain, P. M., Cooper, A., Emmett, B. A., Maskell, L. C., McCann, T., Murphy, J., Norton, L. R., Reynolds, B., Scott, W. A., Simpson, I. C., Smart, S. M., and Ullyett, J. M.: Countryside Survey: UK Results from 2007, NERC/Centre for Ecology & Hydrology, Lancaster, 2008.

Emmett, B. E. and GMEP team: Glastir Monitoring & Evaluation Programme. First Year Annual Report to Welsh Government (Contract reference: C147/2010/11), NERC/Centre for Ecology & Hydrology, Bangor, 442 pp., 2014.

Gimingham, C.: The status of pinewoods in British ecosystems, Native Pinewoods of Scotland, Institute of Terrestrial Ecology, Cambridge, UK, 1–4, 1977.

Goodier, R. and Bunce, R. G. H.: The native pinewoods of Scotland: the current state of the resource, in: Native pinewoods of Scotland, edited by: Bunce, R. G. H. and Jeffers, J. N. R., Institute of Terrestrial Ecology, Cambridge, 78–87, 1977.

Hill, M. O., Bunce, R. G. H., and Shaw, M. W.: Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland, J. Ecol., 63, 597–613, doi:10.2307/2258738, 1975.

ISO 19115 (International Organization for Standardization): available at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020 (last access: 12 October 2015), 2003.

JNCC: 91C0 Caledonian forest, available at: http://jncc.defra.gov.uk/ProtectedSites/SACselection/habitat.asp?FeatureIntCode=H91C0, last access: 2 November 2015.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., and Smithers, R. J.: Long term ecological change in British woodland (1971–2001). A re-survey and analysis of change based on the 103 sites in the Nature Conservancy 'Bunce 1971' woodland survey. Final report, Peterborough: English Nature (English Nature Research Reports Number 653), 139 + appendices, 2005.

Lowe, V. P. W.: Pinewoods as habitats for mammals, in: Native Pinewoods of Scotland, Proceedings of Aviemore Symposium, Insititute of Terrestrial Ecology, Cambridge, 103–111, 1977.

Mason, W., Connolly, T., Pommerening, A., and Edwards, C.: Spatial structure of semi-natural and plantation stands of Scots pine (*Pinus sylvestris* L.) in northern Scotland, Forestry, 80, 567–586, doi:10.1093/forestry/cpm038, 2007.

McHaffie, H., Legg, C. J., Worrell, R., Cowie, N., and Amphlett, A.: Scots pine growing on forested mires in Abernethy Forest, Strathspey, Scotland, Transactions and Proceedings of the Botanical Society of Edinburgh and Botanical Society of Edinburgh Transactions, 54, 209–219, doi:10.1080/03746600208685038, 2002.

McVean, D. N. and Ratcliffe, D. A.: Plant communities of the Scottish Highlands. A study of Scottish mountain, moorland and forest vegetation, Monographs of the Nature Conservancy, HMSO, London, 445 pp., 1962.

Met Office: available at: http://www.metoffice.gov.uk/, last access: 8 April 2016.

Milner, C.: Shetland project monitoring report Nature Conservancy Council, Grange over Sands, Contract Report to the Nature Conservancy Council, Nature Conservancy Council, Grange over Sands, 1975.

Newton, I. and Moss, D.: Breeding birds of Scottish pinewoods, in: Native pinewoods of Scotland, Institute of Terrestrial Ecology, Cambridge, 26–34, 1977.

Palmer, S. C. F. and Truscott, A.-M.: Browsing by deer on naturally regenerating Scots pine (*Pinus sylvestris* L.) and its effects on sapling growth, Forest Ecology and Management, 182, 31–47, doi:10.1016/S0378-1127(03)00026-4, 2003.

Patterson, G., Nelson, D., Robertson, P., and Tullis, J.: Scotland's Native Woodlands: Results from the Native Woodland Survey of Scotland, Forestry Commission, Edinburgh, 2014.

Peterken, G. F.: Natural woodland: ecology and conservation in northern temperate regions, Cambridge University Press, Cambridge, 1996.

Prosser, M. and Wallace, H.: Countryside Survey 1990 Quality Assurance Exercise. Report to Institute of Terrestrial Ecology, Merlewood Research Station, Canon Pyon, Herefordshire, 1992.

Roberts, A., Russell, C., Walker, G., and Kirby, K.: Regional variation in the origin, extent and composition of Scottish woodland, Transactions and Proceedings of the Botanical Society of Edinburgh and Botanical Society of Edinburgh Transactions, 46, 167–189, doi:10.1080/03746600508684786, 1992.

Roche, P. and Geijzendorffer, I.: EBONE: integrated figures of habitat and biodiversity indicators: quantifying indicators of an integrated biodiversity observation system, Alterra, Wageningen UR, 2013.

Romão, C.: Interpretation manual of European Union habitats, version EUR 28, available at: http://ec.europa.eu/environment/nature/legislation/habitatsdirective/docs/Int_Manual_EU28.pdf (last access: 3 May 2016), 2013.

Scott, D., Welch, D., Thurlow, M., and Elston, D. A.: Regeneration of Pinus sylvestris in a natural pinewood in NE Scotland following reduction in grazing by Cervus elaphus, Forest Ecol. Manage., 130, 199–211, doi:10.1016/S0378-1127(99)00191-7, 2000.

Scottish Government: Scottish Biodiversity List, available at: http://www.gov.scot/Topics/Environment/Wildlife-Habitats/16118/Biodiversitylist/SBL, last access: 2 November 2015.

Shaw, M. W. and Bunce, R. G. H.: National Woodlands Classification 1971 Handbook of Field Methods, Merlewood Research Station, The Nature Conservancy, Grange-over-Sands, 1971.

Smith, R.: Botanical survey of Scotland. II. North Perthshire district, The Scottish Geographical Magazine, 16, 441–467, doi:10.1080/00369220008733178, 1900.

Smout, T. C.: The Pinewoods and human use, 1600–1900, Forestry, 79, 341–349, doi:10.1093/forestry/cpl021, 2006.

Steven, H. M. and Carlisle, A.: The native pinewoods of Scotland, Oliver & Boyd, Edinburgh, 1959.

Summers, R. W. and Buckland, S. T.: A first survey of the global population size and distribution of the Scottish Crossbill Loxia scotica, Bird Conservation International, 21, 186–198, doi:10.1017/S0959270909990323, 2011.

Summers, R. W., Proctor, R., Raistrick, P., and Taylor, S.: The structure of Abernethy Forest, Strathspey, Scotland, Transactions and Proceedings of the Botanical Society of Edinburgh and Botanical Society of Edinburgh Transactions, 49, 39–55, doi:10.1080/03746609708684851, 1997.

UK GEMINI: available at: http://www.agi.org.uk/agi-group/standards-committee/uk-gemini, last access: 3 May 2016.

Vickers, A. and Palmer, S.: The influence of canopy cover and other factors upon the regeneration of Scots pine and its associated ground flora within Glen Tanar National Nature Reserve, Forestry, 73, 37–49, doi:10.1093/forestry/73.1.37, 2000.

Wilkinson, N. I., Langston, R. H., Gregory, R. D., Gibbons, D. W., and Marquiss, M.: Capercaillie Tetrao urogallus abundance and habitat use in Scotland, in winter 1998–99: A second full survey in which we estimate the population at 1073 birds (95 % CL 549–2041), Bird Study, 49, 177–185, doi:10.1080/00063650209461263, 2002.

Wilson, B. and Puri, G.: A comparison of pinewood and moorland soils in the Abernethy Forest Reserve, Scotland, Global Ecol. Biogeogr., 10, 291–303, doi:10.1046/j.1466-822X.2001.00226.x, 2001.

Wood, C. M., Caffrey, D., and Bunce, R. G. H.: Dataset documentation – Scottish Pinewoods Survey 1971 (Native Pinewood Survey), Centre for Ecology and Hydrology, Lancaster, 2015a.

Wood, C. M., Smart, S. M., and Bunce, R. G. H.: Woodland Survey of Great Britain 1971–2001, Earth Syst. Sci. Data, 7, 203–214, doi:10.5194/essd-7-203-2015, 2015b.

Earth Syst. Sci. Data, 8, 89–103, 2016
www.earth-syst-sci-data.net/8/89/2016/
doi:10.5194/essd-8-89-2016

Open Access Earth System Science Data

Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring

Claire M. Wood[1] **and Robert G. H. Bunce**[2]

[1]Centre for Ecology and Hydrology, Lancaster Environment Centre, Bailrigg, Lancaster, LA1 4AP, UK
[2]Estonian University of Life Sciences, Kreuzwaldi 5, 51014 Tartu, Estonia

Correspondence to: Claire M. Wood (clamw@ceh.ac.uk)

Received: 12 October 2015 – Published in Earth Syst. Sci. Data Discuss.: 23 October 2015
Revised: 3 February 2016 – Accepted: 14 February 2016 – Published: 25 February 2016

Abstract. A survey of the natural environment was undertaken in Shetland in 1974, after concern was expressed that large-scale development from the new oil industry could threaten the natural features of the islands. A framework was constructed by the Institute of Terrestrial Ecology on which to select samples for the survey. The vegetation and habitat data that were collected, along with the sampling framework, have recently been made public via the following doi:10.5285/06fc0b8c-cc4a-4ea8-b4be-f8bd7ee25342 (Terrestrial habitat, vegetation and soil data from Shetland, 1974) and doi:10.5285/f1b3179e-b446-473d-a5fb-4166668da146 (Land Classification of Shetland 1974). In addition to providing valuable information about the state of the natural environment of Shetland, the repeatable and statistically robust methods developed in the survey were used to underpin the Countryside Survey, Great Britain's national long-term integrated environmental monitoring programme. The demonstration of the effectiveness of the methodology indicates that a repeat of the Shetland survey would yield statistics about ecological changes in the islands, such as those arising from the impacts of the oil industry, a range of socio-economic impacts, and perhaps climate change. Currently no such figures are available, although there is much information on the sociological impacts, as well as changes in agriculture.

1 Introduction

In the 1960s, the discovery of North Sea oil off the coast of Shetland, Scotland, meant that the islands had to face the prospect of large-scale development to accommodate the infrastructure surrounding the industry. In the early 1970s, concern was expressed by the county council that such development would threaten the natural features of Shetland, such as landscape and wildlife. Accordingly, in 1974, a survey was commissioned by the Nature Conservancy Council (now Scottish Natural Heritage in Scotland) and organised by the Institute of Terrestrial Ecology (ITE) (now part of the Centre for Ecology and Hydrology) to assess the natural environment of the islands. Although the terrestrial habitats and vegetation survey are the focus of this paper, assessments of freshwater and littoral habitats, sea-bird populations, geology and geomorphology were also undertaken around the same time as part of a larger integrated survey (Milner, 1975).

The specific objectives of the terrestrial component were as follows:

i. To assess the range of variation within the vegetation of Shetland and to provide a user guide to defined types.

ii. To provide a structural basis for monitoring future change in the vegetation (Bunce, 1975).

The Shetland Survey was also a stage in the development by ITE of the methodology of strategic ecological survey, as described by Sheail and Bunce (2003).

The vegetation survey was undertaken using standardised methods outlined by Bunce and Shaw (Bunce and Shaw, 1973; Bunce, 1974). In order to randomly sample Shetland in a strategic way, the islands were stratified into a set of 16 relatively homogenous areas (strata, or "land classes"). Sampling locations were randomly selected from within each of these strata, giving a total of eighty 1 km^2 sampling units, each containing up to 16 200 m^2 sampling plots. Records of plant species, soils, habitat types and major biota present were collected.

Prior to the survey, a two-day training course was held to familiarise the surveyors with the detailed field protocols. Additionally, all survey teams were initially accompanied by a supervisor, and regular visits into the field were made by the project leader to ensure consistency and quality in data recording according to criteria laid out in the field handbook (Bunce, 1974).

At the time, the statistical sampling methods used were generally not widely used as a method for ecological monitoring, and this was one reason why the results were not reported publicly at the time. The other factor was that the senior author also had other commitments. Therefore, this is the first time that the data and results have been made widely available. There have been many changes in the islands since the survey, both socio-economic and environmental, and the data provide a unique opportunity to explore the changes that have taken place in the vegetation, should a repeat survey be undertaken.

In addition to the survey yielding an interesting set of data in itself, the methodological framework for the survey eventually developed into the largest long-term ecological monitoring project in Britain, the Countryside Survey (CS) (Carey et al., 2008). CS started in 1978 and was most recently undertaken in 2007.

2 Shetland

The Shetland Islands cover an area of about 1400 km^2 and consist of over a hundred islands and islets, of which about 15 are inhabited. The southernmost tip of the largest island lies over 160 km north-west of John o' Groats. Lying at the northern limit of the Britain, the isles have considerable biogeographic interest.

The islands are geologically diverse, with the main rock types being metamorphic, including Caledonian schists, gneisses and quartzites. There are also areas of Old Red Sandstone.

The inland topography is gentle, with wide, shallow valleys. Around the coast, there are sheer cliffs, and numerous sheltered inlets or "voes". There are few trees on the islands, and extensive areas are covered in peat, especially on the Mainland and the northern island of Yell. The soils on Shetland are generally poor, with the most fertile land being on the sandstone in the south, where the main crofting districts are located.

The general climate is mild, moist and windy. Although minor changes may be discerned over time (see Sect. 3.3), there is little variation in temperature through the year, with the average monthly temperature ranging from approximately 3 °C (February) to 12 °C (July) (Met Office, 2015). The exposed situation of the islands means they are subject to high winds, with about 40 days of gales per year. The rainfall is not extreme (around 1124 mm per annum) (Met Office, 2015) but is distributed throughout the year, so that damp and drizzly days are common.

3 The survey in context

The survey of Shetland is relevant on a number of levels. Firstly, the data are of local interest to anyone concerned with the ecology and land use of the islands. Should a resurvey be undertaken, potential changes in vegetation could be assessed against a range of local and national socio-economic changes occurring since the 1970s. Secondly, the data are of a wider global significance when applied to the potential effects of global climate change.

Additionally, as described in Sect. 4, the design of the methodology itself is of importance, having been applied to national monitoring schemes in Britain, and beyond.

3.1 The survey in a local context

At the time of the survey, a review of the available ecological knowledge regarding Shetland was made by Goode (1974). Attention was drawn to the major gaps that were present, few of which have since been filled. Insufficient information and detail were available on the islands in McVean and Ratcliffe (1962). Birse (1974) provided a general account of Sullom Voe and Barkham (1971) and Allot (1971) described aspects of Foula. The specialised habitats of the fell fields and the serpentine are also summarised by Spence (1974). Most of the published work has concentrated on two specialised habitats – the Fell fields of Ronas Hill and the serpentine habitats of Unst (Spence, 1974). The small areas of relict scrub, mainly on ungrazed islands in the lochs, are also described by the same author.

The only major paper on the overall vegetation is by Roper-Lindsay and Say (1986), who used phytosociological methods to describe 17 associations in relation to British, continental and Scandinavian communities. They found difficulties in determining discrete associations because of factors such as intensive land use and the maritime influence. Hence, the present study is the first to provide a complete overview of the vegetation of the islands.

By contrast, the flora of Shetland is relatively well known, for example Scott and Palmer (1987) for vascular plants and Dalby and Dalby (2005) for lichens and as summarised in the BSBI Atlas of the British and Irish Flora (Preston et al., 2002), which is due to be updated in 2020, and the original bryophyte atlas published in 1991–1994 (Hill et al., 1991) but recently repeated (Blockeel et al., 2014). Also, Hill and Paton (1976) have reported on the saxicolous bryophytes. The phyto-geographic relationships of the Shetland flora have been widely discussed, as summarised by Goode (1974). Although some species differ in their ecology because of the northern location, the species complement is closely related to that of northern Scotland. Recent overviews of the ecology of the islands are given by Berry and Johnston (1980) and Johnston (1999).

Superficially, the vegetation is similar to northern Scotland except for the extensive maritime influence and the distinctive vegetation on the serpentine, which is similar to that of comparable outcrops elsewhere, for example on Rhum in western Scotland. Although at a lower altitude than in northern Scotland, the sub-arctic vegetation on Ronas Hill is otherwise similar to high altitudes on other Scottish mountains.

In order to obtain a general comparison with the environment and land cover of Shetland with the rest of Britain, data from Britain's national monitoring programme, the Countryside Survey (Brown et al., 2014), may be used. The overall land cover, described by the CS survey data, is dominated by bogs and acid grassland habitats and provides an independent comparison with the results from the vegetation survey described below. In terms of a general comparison of the environmental characteristics of Shetland with the Outer Hebrides and Orkney using the environmental classes of the ITE Land Classification (Bunce et al., 2007), the former is shown to indeed be an outlier and is dominated by the most pronounced overall land class – 32. Class 32 can be summarised as having variable topography, mainly at medium/low altitudes, peaty soils, peatland and moorland vegetation types, and a landscape of scattered lochs and eroding peat hags (Benefield and Bunce, 1982). All the other classes in the islands, except one, a marginal upland class, are also upland, showing that Shetland is upland in its affinities even though it has no land at high elevations.

3.2 Drivers of potential change – socio-economic factors

Overall, Shetland has undergone a range of changes since the 1970s, largely driven by economic and political factors, most notably the new oil industry. Wills (1991) provides a thorough account of the overall impact of the new Shetland oil industry, describing how the new multi-million-pound terminal at Sullom Voe had major impacts on the traditional culture, economy and environment of the islands. The conflict between economic growth and biodiversity conservation has often been discussed (for example in Chambers et al., 2014; Czech, 2003), and in the case of smaller island communities, the introduction of economic growth tends to have an even greater impact on the environment than might be the case elsewhere (Daly and Farley, 2011). Examples of this have been explored in islands around the world, for example in Sakhalin, Russia (Wilson, 2003), and most recently in the Caribbean (Huettmann, 2015).

In Shetland, there has been a general decline in traditional crofting agriculture, partly due to the availability of well-paid jobs in the oil industry, and factors such as EU subsidies. Key habitats and species of conservation value are often found in crofting communities; therefore these are now under threat (Scottish Rural Development Programme, 2016). Additionally, the demography of the crofting community is changing and many common grazings now have insufficient active adults to undertake routine operations such as sheep round-ups. This has seen fairly widespread apportionment (fencing of hill shares) which has broken up large hill units.

Sheep numbers have fluctuated in response to headage subsidies from the EU, from around 265 000 in 1971 to nearly 400 000 in 2001 (Scottish Government, 2016), with the associated grazing pressures. This figure reduced to around 280 000 in 2010 as a result of subsidies switching to area payments in the early 2000s and agri-environment schemes encouraging better stewardship of the land. Additionally, the amount of land put to silage has increased markedly since the 1970s, whilst at the same time the amount of arable has dropped (Scottish Government, 2016),

The availability of grant aid from the Local Authority (as a result of oil income) meant that there was a considerable amount of re-seeding, surface seeding or liming of moorland/heathland, as well as large drainage schemes. Now that the economic situation is currently less favourable, there is no money available to crofters for these activities, and much of this improved land is reverting back (Scottish Natural Heritage, 2002).

Peat is frequently used as a domestic fuel in Shetland. Peat cutting has declined, although there has been a small resurgence in recent times, in response to high oil prices. Mechanised cutting can seriously damage blanket bog by draining the peat and destroying large areas of surface vegetation. However, there are now only a few commercial peat cutting operators and the practice is currently viewed as unsustainable (Scottish Natural Heritage, 2002).

The construction of the third largest wind farm in Scotland (Viking Energy, 2016), due to be built in the near future, could also affect the vegetation of this islands by altering the local microclimate.

Sulfur deposition from shipping is also likely to have had a certain impact on the Shetland environment, as well as oil spillages. A notable oil spillage occurred in 1993, when the MV *Braer* ran aground off the islands, carrying 85 000 tonnes of crude oil. The significant and persistent environmental effects of major oil spills such as these is emphasised by Ott (2005) in relation to the 1989 *Exxon Valdez* spill off Alaska.

3.3 Drivers of potential change – climate change

It would be expected that the most significant drivers of changes in the vegetation and habitats would be largely the socio-economic factors described. However, there is also the possibility that climate change may have already had some effect on the Shetland vegetation, and may do so in the future. Changes in the climate since the 1970s have not been great,

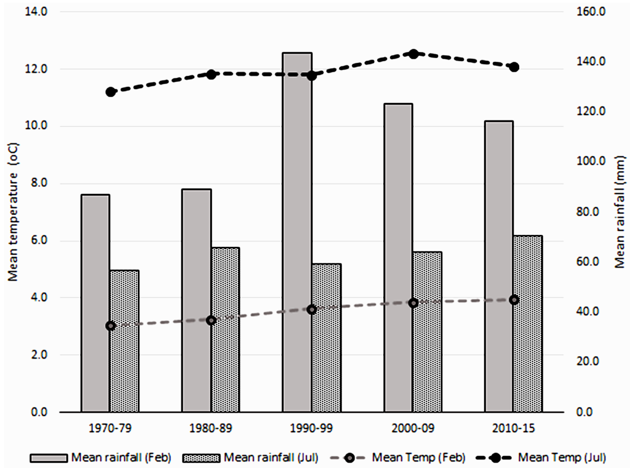


Figure 1. Climate trends in Shetland 1970–2015 (Met Office, 2015).

but from the climate data presented in Fig. 1, there appears to be a trend towards higher temperatures and wetter winters. In the next 50 years, average temperatures are predicted to rise by 0.5–1.5 °C, with 6–13 % more rainfall (Scottish Natural Heritage, 2002).

In a European context, Shetland is located in the Atlantic North zone, as described by Metzger et al. (2008). The Atlantic environment is relatively stable compared to other regions, and is not expected to change dramatically in extent or location in comparison to other zones (Metzger et al., 2008). However, the effects of potential climate change are difficult to predict. The species and habitats most likely to be affected are those which are close to the limits of their range, such as arctic–alpine plants. In Shetland, due to the cool winds, arctic–alpine species are present at much lower altitudes than would be expected on the British mainland. These species include *Carex bigelowii* and *Silene acaulis*. A warmer climate would perhaps result in a loss of these arctic–alpines and an increase in more generalist species, such as *Agrostis tenuis* and *Festuca rubra*. The Global Observation Research Initiative in Alpine Environments (GLORIA) project has yielded research from the Swiss, Austrian and Italian Alps showing that warming causes arctic–alpine species to retreat to elevations higher than those pre-warming (Grabherr et al., 2010). As Shetland does not have higher elevations, it is likely the arctic–alpine species there would be lost in the event of climate warming. Monitoring these species in the islands would provide an early warning of the effects of climate change. Additionally, as this survey covers a range of functional plant types, changes in these could also be analysed in the context of climate change.

4 Survey design: site selection and stratification

The method used to undertake the survey initially requires that an environmental classification of the whole area in question (in this case, Shetland) is constructed using multivariate analysis of environmental characteristics to produce a set of strata, or areas of relatively homogenous regions. These strata can then be used for randomly selecting sample sites in order to sample ecological parameters such as vegetation. By using this statistically robust method, it is then possible to scale up the results from the sample sites to describe the entire population.

By the end of the 1970s, this idea of stratifying the landscape led to the successful creation of a stratification for the whole of Great Britain, known as the "Institute of Terrestrial Ecology (ITE) Land Classification of Great Britain" (Bunce et al., 1990, 1996a, b). Although this has developed over time (Bunce et al., 1998, 2007), the basic stratification still underpins the CS (Carey et al., 2008).

4.1 Land stratification – "ITE Land Classification of Shetland"

This "ITE Land Classification of Shetland" (Bunce and Bassett, 2015) separated the 2046 km^2 of the grid of the national mapping agency, the Ordnance Survey, into 16 relatively homogeneous units (known as strata, or "land classes") with similar environmental characteristics and geographical features, ascertained from a map study. Most of the criteria were derived from one-inch-to-a-mile (1 : 63 360) Ordnance Survey maps, with the 18 geological attributes being recorded from the quarter-inch geological map (British Geological Survey, 1963), which was the only one to give complete cover. It was decided to use the 1 km grid squares as the sampling unit, as these have the advantage of being fixed and readily referable between maps of different scales.

The range of geographical factors on a map can be broadly divided into physical attributes (such as hills, valleys, coastlines) and features of human geography (for example roads and houses), and it was decided from the outset to consider only the physical geography of Shetland in order to provide a classification which would be readily interpretable in terms of landform.

The geographical features fall into two types:

i. Continuous, e.g. altitude and slope, which can be represented as an integer or decimal number.

ii. Attribute data, in which the feature is either present or absent, e.g. a cliff or hill top.

The full list of 150 attributes is given in supporting documentation supplied with the data set (Bunce, 1975).

There are 2046 one-kilometre squares in Shetland containing some land, which were each allocated to one of the 16 strata using indicator species analysis (now TWINSPAN)

Table 1. Description of strata.

Group	Strata	Description
1	Strata 1–4	Coastal strata with few rivers running into the sea within the square. There is more than 80 % land area, and terrain is relatively gentle.
2	Strata 5–8	Coastal group with more sea and steeper slopes. It is more likely to contain headlands and sea cliffs. There are also more likely to be more rivers entering the sea.
3	Strata 9–12	High altitude inland group with a 600–900 ft (182–274 m) hill within the square or close by. There are few small water bodies and the major rock is likely to be gneiss.
4	Strata 13–16	Lower, more undulating group with much peat and many freshwater lochans. The hills are about 300 ft (91 m) and the rock is more likely to be Old Red Sandstone.

described in Hill (1973) and Hill et al. (1975) as shown in Figs. 2 and 3. The environmental classes were produced using an arbitrary stopping rule. The first division of the environmental classification was on the basis of inland versus coastal features (Fig. 3). The eight coastal types are distinguished by the different proportions of sea and inlets such as voes. The eight inland classes are based on exposure and drainage patterns. Overall, it was found that the square's relationship with the sea dominated the classification. A summary description of the classes is given in Table 1.

4.2 Sampling sites and plots

The sample survey sites were based on 1 km^2 units randomly selected from across Shetland, as shown in Fig. 2. Five of these square units were randomly chosen from each of the 16 strata (or land classes) for survey, giving a total of 80 1 km^2 units to be surveyed.

Within each selected 1 km^2 unit (or survey site), 16 individual 200 m^2 plots were chosen for detailed survey (Fig. 4). In order to allocate the location of these, each of the 80 selected 1 km^2 units was split into 16 equal sub-squares using an overlay grid, and a point location for the 200 m^2 plot was then randomly selected from within each of the 16 sub-squares. If a sub-square contained no land, that sub-square (and hence plot) was omitted from the survey; therefore, between 1 and 16 plots were surveyed for each of the 1 km^2 units. In total, 927 plots of 200 m^2 were selected, although only 911 of these plots were actually surveyed as occasionally it was not possible to record the plot due to inaccessibility of a particular piece of land, for example due to a dangerous cliff or cultivated land. The location of each plot was marked on a one-inch-to-a-mile (1 : 63 360) map.

A key aim of the sampling design was that the methods chosen should be standardised, therefore highly repeatable. The size of the plot was chosen with reference to continental phytosociologists, who at the time most widely used plots of between 100 and 200 m^2 (Bunce and Shaw, 1973). After preliminary field tests, it was found that the number of species recorded usually stabilised at this size. The area

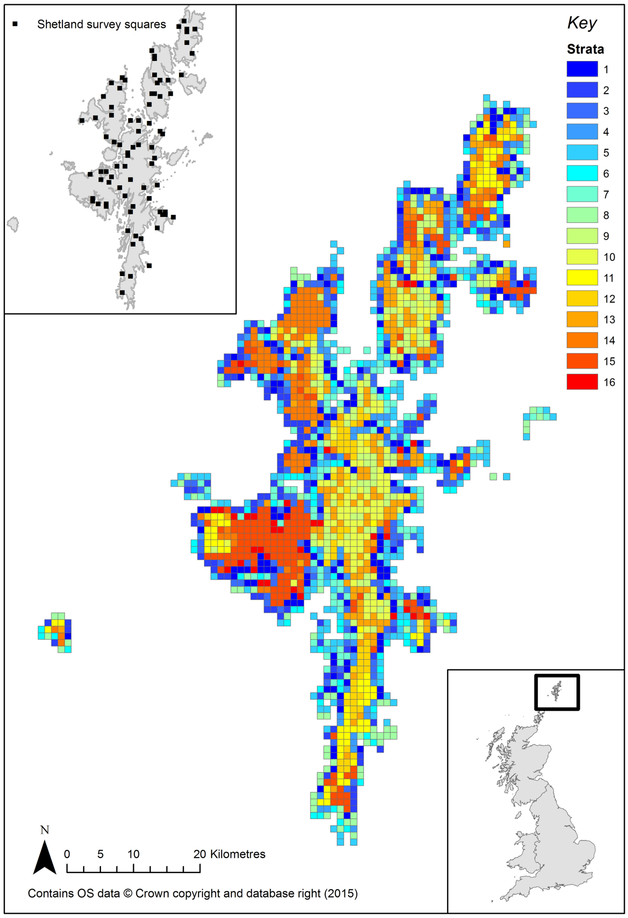


Figure 2. Map of sampling sites and strata.

of 200 m^2 was thus adopted for this survey, with five nests within (Fig. 5). This design of plot aids a systematic search of the vegetation within; it is also straightforward to layout in the field, and ensures a standard-sized plot is laid out every time. The plots were constructed as shown in Fig. 6, with one centre post and four corner posts, with a set of four strings tagged with markers at specified distances. The centre post

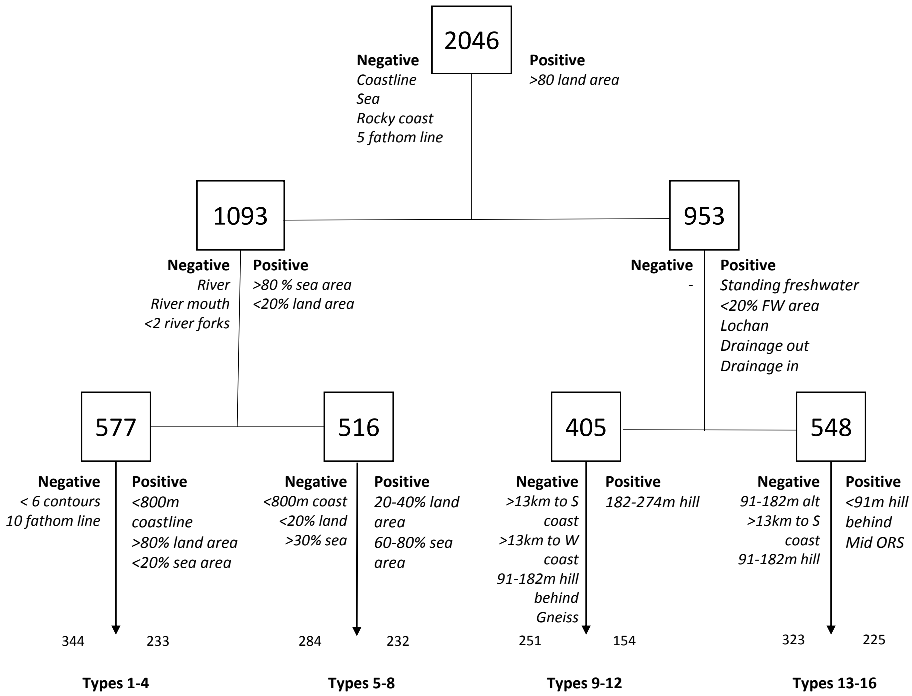


Figure 3. Hierarchy of divisions in the land classification (numbers refer to number of 1 m grid squares).

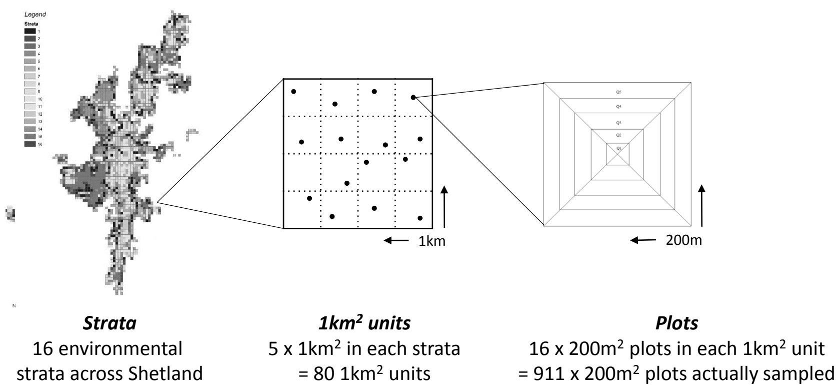


Figure 4. Sampling design.

had a right-angled gauge affixed to the top in order to orientate the plot at random.

In the field, plots were located using the prepared map, and then by pacing from the nearest relocatable feature. It was stressed in the field handbook that there should be no subjective bias in locating the plots on the ground, and the plots in the field should be located as accurately as possible from the map. Data were then collected on ground flora, soils and habitat characteristics for the plot. A habitat sheet for the area within 50 m of the plot was also compiled.

The sampling intensity of 16 plots per site was originally used in a previous survey, the Woodland Survey of Great Britain in 1971 (Wood et al., 2015), and was chosen on the basis of previous experiences in surveying a wide variety of sites in the north of England and Wales. It also coincided with the time and manpower available (Bunce and Shaw, 1973).

5 Data collected

An overview of all the data collected from the survey is given in Table 2. These categories are described in more detail in the following subsections.

5.1 Site information, plot locations and information, slope and aspect

For each plot surveyed, both from within the 200 m^2 plot and from the surrounding area to within 50 m of the edge

Table 2. Summary of data collected.

Data category	Description
Ground flora	Vascular plants, bryophytes and lichens present in the plot % cover/abundance estimates
Plot description and habitats	Tick list of features (broad categories): – Rock habitats – Aquatic habitats – Open habitats – Vegetation structure – Animal signs – Management – Land use
Soil data	Tick list description from small pit and augur boring in the centre of the plot – to determine soil type Composite soil sample from top 10–15 cm taken for pH
Within 50 m of plot description	Tick list of features (broad categories) – as for plot, plus adjacent land use and boundary type Slope, aspect

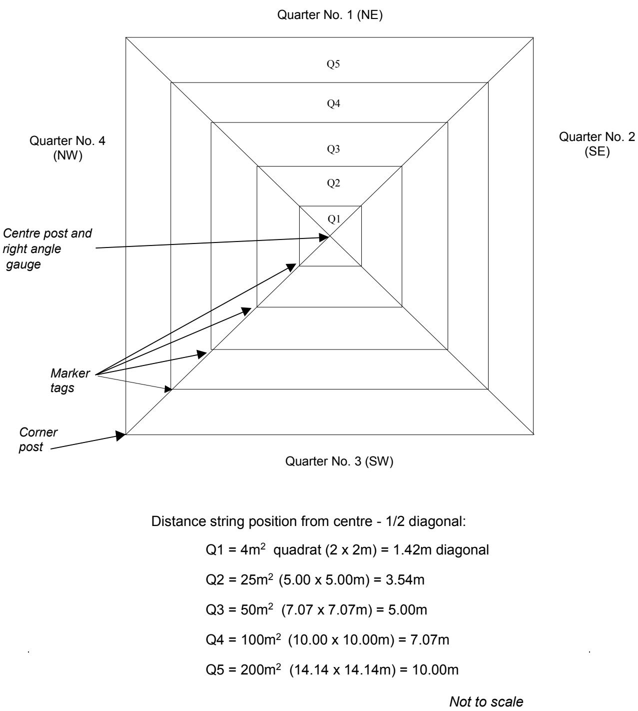


Figure 5. Layout of vegetation plot.

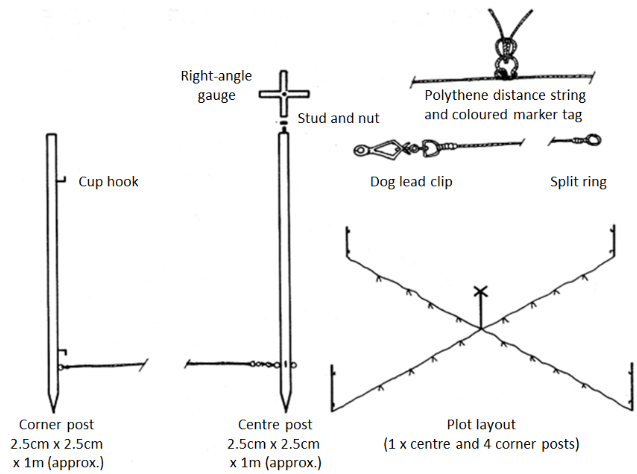


Figure 6. Plot construction.

of the plot, the presence and absence of a series of attributes were recorded. Attributes included physical factors such as the presence of rock or cliffs; habitat-related factors such as the presence of heather, bryophytes and trees; aquatic habitats such as ponds; presence of buildings, quarries or rubbish; presence of animals and birds; and also boundary types and nearby land use. All classes within each of the pre-defined habitat categories which were relevant to the survey area were marked off on the appropriate data sheet. Precise definitions for each habitat category and its classes were provided, and a full list of habitats may be found in the 1974 field handbook (Bunce, 1974) (supplied as supporting documentation with the data sets). The slope of each plot was measured using a clinometer from the highest to lowest point passing through the centre of the plot. The aspect was taken bearing down the slope, measured with a magnetic compass.

5.2 Vegetation data

Within the plot described in Fig. 5, the area within the first nest of the plot (2×2 m) was searched for the presence of all vascular plants (monocots, dicots, gymnosperms and ferns, including tree species) as well as bryophytes and macro-lichens growing on soil. This procedure was repeated for each nest of the quadrat, increasing the size each time as shown in Fig. 5. In the final nest (the whole 200 m^2 plot), the percentage cover (to the nearest 5 %) of each species was estimated. In addition, the total cover of bryophytes was estimated from the entire plot, as was an overall estimate for litter, wood, rock, bare ground and standing water.

5.3 Soil data

The soil of each plot surveyed was classified by horizon using a set of standard categories. In the centre of each plot a shallow pit was dug to enable examination of the surface layers of soil, and auger samples were taken to classify lower horizons. Precise definitions for each of the descriptive categories used and are detailed in the field handbook. A sample from the top 10 cm was taken away for pH analysis. A single composite soil sample was taken from each plot, at the centre of the vegetation quadrat, using a trowel. Samples (weighing approximately 1 kg) were taken to a depth of 15 cm and placed in a labelled plastic bag. A pH reading was taken on a representative fresh subsample from each soil sample, using a field pH meter.

6 Data quality and repeatability

During the survey, all survey teams were initially accompanied by a supervisor and regular visits into the field were made by the project leader to ensure consistency and quality in data recording according to criteria laid out in the field handbook (Bunce, 1974).

The data sets were transferred from the original field sheets to spreadsheets in the 2000s. They were checked and corrected to produce a final validated copy. Standard validation checks included plot and site counts to ensure no duplicate numbering and hence double counting of plots; also, range checks were undertaken where possible for values falling within certain ranges, such as soil pH or slope values.

In terms of the soil data, the descriptive profile data were collected to the standards set out in the training and field handbook, but these were not formally checked for quality aside from checks from supervisors during the survey.

In terms of the representativeness of the plot information across a wider area, an analysis is presented in the survey report (Bunce, 1975). This analysis was undertaken on the Walls Peninsula, an area of Shetland, and compares the results gained from the plot information against information from aerial photographs. The correspondence between the different values was found to be reasonably close at a broad vegetation category level (Table 3).

Table 3. Comparison of broad vegetation group covers on the Walls Peninsula: plot estimates versus aerial photography estimates.

Group	Plot estimate of cover %	Aerial photo estimate of cover %
Calluna/Eriophorum	29.5	39.0
Calluna/Nardus	23.0	25.0
Nardus	14.9	17.0
Juncus effusus	10.1	3.0
Calluna/Rhacomitrium	5.4	10.0
Festuca rubra	4.7	3.0
Agrostis/Holcus	11.4	6.0

The methodology has subsequently been developed further (for example in CS; Maskell et al., 2008b) to include a habitat mapping component to capture the variation in habitats across 1 km squares. This could perhaps be usefully incorporated into a repeat survey of Shetland.

In terms of repeatability, it is certainly the case that statistical analyses of temporal vegetation change are more powerful when based on records from plots located in the same place rather than randomised to new locations at each survey. Surveys using exactly the same methodology as in Shetland have been proven to be highly repeatable. One such example is the Woodland Survey of Great Britain, carried out in 1971 and again in 2001. In the repeat survey, the field surveyor relied only on the marked point on a map as the sole aid to relocating the 1971 plot location (as would be the case in a repeat survey of Shetland). Consequently, there is the potential for considerable relocation error. The expectation is that, having made an effort to move near to the mapped point, the plot records from the repeat survey will, on average, be more similar to the respective 1971 plot record than if a completely new, random set of locations were chosen. Even if vegetation change occurs, species compositional data recorded from the same point at times 1 and 2 will tend to be more similar than data recorded from two random points at times 1 and 2. In attempting to measure the amount of relocation error, one cannot of course exploit a "true" set of temporal pairs known to have been recorded in exactly the same position. What can be done is to compare the average species compositional similarity between the ostensibly true temporal pairs with the average similarity for a random pairing of the 1971 data with the 2001 data. If, on average, attempts to relocate the true 1971 position had been successful, then the similarity between the true pairs should be greater than the random pairs. Overall, at 97 of the woodland sites (out of 103), mean similarity was greater between "relocated" plot pairs compared to random-pair comparison; for 59 sites the difference was significantly greater – therefore we were satisfied that the relocation error was not significant when interpreting any results. A full account of this is given in Appendix 3 of Kirby et al. (2005).

7 Analysis to date

Aside from the initial analyses undertaken in the final report in 1975, little analysis has been undertaken using the data since, largely due to the authors' other commitments and the inaccessibility of the data set to other workers. The analyses undertaken to date are described in the following subsections.

Table 4. List of the 25 most abundant species recorded in the survey.

	Species	Total records
1	*Hypnum cupressiforme*	741
2	*Calluna vulgaris*	740
3	*Potentilla erecta*	699
4	*Dicranum scoparium*	681
5	*Carex panicea*	648
6	*Nardus stricta*	638
7	*Rhytidiadelphus loreus*	608
8	*Cladonia impexa*	607
9	*Juncus squarrosus*	601
10	*Sphagnum rubellum*	598
11	*Eriophorum angustifolium*	596
12	*Cladonia uncialis*	550
13	*Rhacomitrium lanuginosum*	547
14	*Scapania sp.*	542
15	*Agrostis canina*	536
16	*Luzula multiflora*	535
17	*Trichophorum caespitosum*	525
18	*Empetrum nigrum*	507
19	*Hylocomium splendens*	487
20	*Mnium hornum*	482
21	*Anthoxanthum odoratum*	447
22	*Festuca vivipara*	434
23	*Erica tetralix*	433
24	*Agrostis tenuis*	433
25	*Campylopus flexuosus*	431

7.1 Vegetation survey

The frequencies of species found in the survey confirm much of what has been written about the phyto-geography of the islands, in that the majority of the species are wide-ranging members of heath and bog communities throughout Scotland. The arctic–alpine species present in Shetland are of restricted distribution. However, those that are present often (atypically) extend to sea level. The dominant species found in the survey are typical of the northern habitats of Britain. Along with several species of mosses, *Calluna vulgaris* had the highest frequency, followed by *Potentilla erecta*, *Carex panicea*, *Nardus stricta* and *Juncus squarrosus* (Table 4). (Nomenclature follows that of Clapham et al., 1952, as used at the time of the survey.)

In the original survey report (Bunce, 1975), indicator species analysis (ISA) (Hill et al., 1975) (now TWINSPAN) (Hill and Šmilauer, 2005) was used to classify the species lists into 32 classes. The first division was on the basis of heathland species such as *Empetrum nigrum* as opposed to grassland species such as *Holcus lanatus*. Subsequent divisions separated classes such as the serpentine grasslands and different types of blanket bogs.

Species groups were constructed from the vegetation data and were designed to help explain details of the structure of the vegetation in Shetland. The groups differed widely from those in Britain and reflect the unique nature of the vegetation of the islands. Details are given in the full report with a discussion of their characteristics (Bunce, 1975). Detailed descriptions of the vegetation classes are included, including the most frequent and dominant species.

The majority of the vegetation is associated with blanket peat, which also intergrades with acid grassland. Agricultural grasslands and serpentine grassland diverge as separate groups. Surprisingly, only two groups are directly controlled by management – peat cutting and burning; the reminder are controlled by environmental factors.

In the 1975 report (Bunce, 1975), the principal axes of variation in the vegetation were correlated with those of the environment using orthogonal regression. The primary axes were correlated with an r value of 0.877, showing that the main gradients are highly correlated. The first axis of the environment is from acid blanket bogs to fertile grasslands, which compares with a gradient of wet peat soils of low pH as compared with mineral moist soils of relatively high pH. The factors that limit peat formation are the relatively nutrient-rich bed rocks, the maritime influence and good drainage. All the fertile soils have been used for agriculture, which has increased inputs such as fertiliser and manure adding to the fertility and further divergence from the peatlands. Further detailed discussion is available in the report (Bunce, 1975), but the conclusion is that, with such high correlation, there would have to be major changes in nutrient inputs to deflect the status of a given patch of vegetation in the short term. However, if agricultural inputs had declined to a major degree in the 40 years since the survey, it would be expected that there would be a shift towards more acidic soil conditions.

7.2 Resource assessment

As the relationship between the vegetation plots and the environmental land classes is known, the coverage of the plot types throughout the islands may be calculated. The 32 plot types described above were grouped into convenient classes and key species identified as follows, presented with an overall estimate of the percentage of the islands covered:

Calluna/Eriophorum 37.1 %, *Calluna/Nardus* 23.3 %,
Nardus 12.7 %, *Juncus effusus* 11.2 %,
Calluna/Rhacomitrium 3.0 %, *Festuca rubra* 4.4 %,
Agrostis/Holcus 8.3 %.

This analysis was the first of its type where cover was calculated from a statistical sample and shows the dominance of bog and heathland vegetation in the islands. By performing analysis in this way, maps may be produced to show the distribution of different vegetation types, an example of which is shown in Fig. 7, showing the estimated distribution of the *Calluna/Eriophorum* group. The basics of this type of analy-

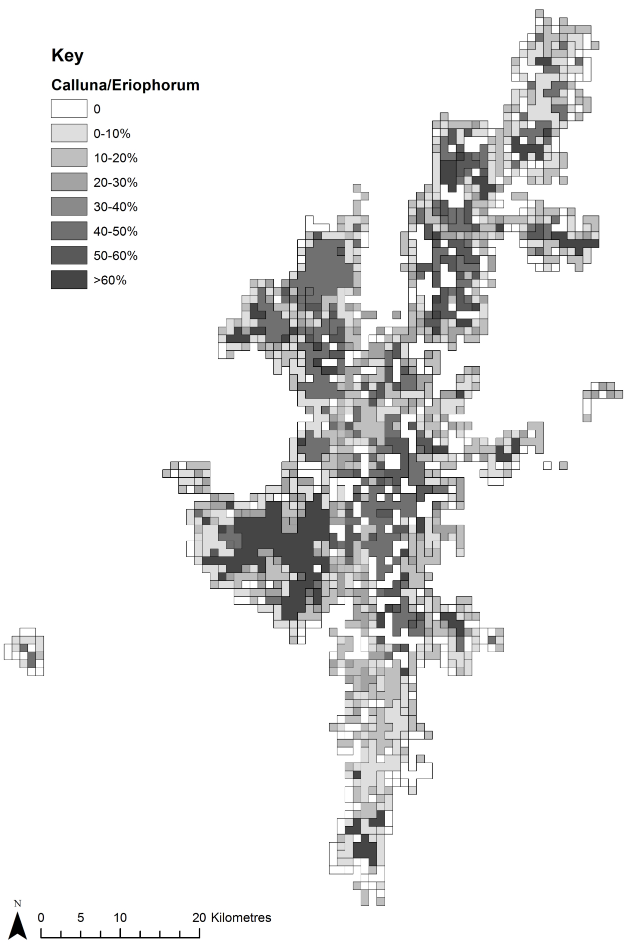


Figure 7. Estimated distribution of the *Calluna/Eriophorum* group.

sis later came to form the basis of calculations of the national estimates of broad habitats (Jackson, 2000) in Britain, as calculated from Countryside Survey data (Scott, 2008; Barr et al., 2014a, b; Bunce et al., 2012; Brown et al., 2014).

8 Methodology as a framework for long-term monitoring

The survey methods were based on those first successfully developed for surveying woodlands across Great Britain (Wood et al., 2015; Bunce and Shaw, 1973). Whereas that survey had focused on woodland sites, the Shetland survey was the first time that samples were being used to obtain an integrated assessment of the response of vegetation to the environment across a defined population across a whole landscape and range of habitats. It was the first project to complete all the stages of land classification, survey, statistical analysis of vegetation and environment through to the estimation of the extent and distribution of ecological resources. The structure of the project provided the basis for the further development of strategic survey methods. The methods developed throughout the 1970s and continued to be tested on a regional basis, as in the Cumbria Survey (Bunce and Smith, 1978).

Although the Shetland survey took place over 40 years ago, the basic methodology has come to underpin much larger and more significant surveys across the whole of Great Britain. The 1 km square unit sampled at random, with random plots sampled within, became a standard sampling strategy, variations of which are currently used very successfully in several large ecological surveys in Britain, such as the Countryside Survey (Carey et al., 2008) and the Glastir Monitoring and Evaluation Programme (Emmett and GMEP team, 2014). In these surveys, the methods are now implemented very successfully using handheld computers to assist recording the field data and global positioning systems to record the location of the vegetation plots. The development of geographical information systems has greatly facilitated the types of analyses able to be undertaken using this kind of data, and has enhanced the use of the capturing ancillary explanatory data sets, as described in the Integrated Assessment Report for the Countryside Survey (Smart et al., 2010). Outwith the UK, within the European Biodiversity Observation Network (EBONE), methods adapted from the basic principles of the Bunce and Shaw methods have been developed to roll out across the whole of Europe (Bunce et al., 2011, 2008). During the EBONE project, the methods were widely tested across 12 European countries, and also Israel, Australia and South Africa. The methods were proven to be robust, reliable and repeatable at a continental, landscape scale (Roche and Geijzendorffer, 2013).

The repeatability of the methods has been proven in large surveys such as the Woodland Survey of 1971, repeated in 2001 (Wood et al., 2015), and the Countryside Survey (CS) of 1978, repeated in 1990, 2000 and 2007 (Carey et al., 2008; Norton et al., 2012; Sheail and Bunce, 2003). The plots from both of those surveys were sufficiently accurately marked on 1 : 10 000 maps in order to relocate the plot again accurately, as discussed in Sect. 6. Changes in vegetation have been monitored in Britain at roughly 10-year intervals since 1978 via the CS, demonstrating the robustness of the procedure. The major difference between the Shetland Survey and the CS being that 16 (200 m^2) plots were surveyed in each kilometre square on Shetland, whereas only 5 (of that size) were used in the CS, partly because of the experience gained in Shetland, and partly due to the time considerations in surveying the whole country. In the CS, although there are fewer plots (5 × 200 m^2 plots), a habitat mapping component has been introduced, whereby areas of broad and priority habitats (Jackson, 2000) are mapped across each 1 km^2, giving the ability to estimate the overall stock of these habitats across the country. Additional smaller plots are also recorded (Maskell et al., 2008a). In Shetland, as in the individual woodlands of the Great Britain Woodland Survey (Kirby et al., 2005; Wood et al., 2015), the data from the 16 plots are relied upon to describe the overall population.

There are many successful environmental monitoring programmes in Britain (for example the Environmental Change Network, UK Environmental Change Network, 2015; the National Soil Inventory, Cranfield University, 2015; the UK Forest Inventory, Forestry Commission, 2013), Europe (for example Long Term Ecological Research (LTER) programmes; Parr et al., 2003) and worldwide (for example Global Earth Observation System of Systems; Lautenbacher, 2006). All of these have their own merits; however, many programmes have a narrow focus, and most have different measurement protocols and sampling designs (Parr et al., 2002). More recently, a consideration of habitat monitoring in Europe by Lengyel et al. (2008a) highlights several factors that would be desirable when planning monitoring schemes but which are factors often lacking. In particular, spatial variations are barely monitored in over half of the monitoring schemes in Europe, schemes pre-1990 are rare, and few schemes use advanced statistical methods to present the data. In terms of the Shetland survey, spatial variation is a key factor in the design of the survey; the survey is historically significant, being first undertaken in 1974; and the data have been collected in such a way that advanced statistical methods can be applied effectively.

When considering the integration of monitoring programmes across geopolitical zones, issues remain such as the ability to aggregate data across scales, the consistent monitoring of global biodiversity change, and linking of in situ and earth observations (Metzger et al., 2013a). The underlying principles of the Shetland/CS methodology provide an ideal framework for the planning of large-scale monitoring, not only in Britain but also across Europe and worldwide. Metzger et al. (2013a) describe how the idea of environmental stratification as a basis for environmental survey would be ideal to start addressing some of these issues. Indeed, stratifications of Europe and the world have already been produced (Metzger et al., 2005, 2013b). The integration of surveys across Europe is discussed in Lengyel et al. (2008b), who recommend methodology such as BioHab (Bunce et al., 2005) in order to increase the potential for integration. By using the general habitat categories described in Bunce et al. (2005, 2008), data from this Shetland survey could be analysed in a European context and could enhance the understanding of landscape ecological change, in terms of a range of drivers.

9 Conclusions

Considering the increased popularity in tourism and the high profile of Shetland in national news since the oil industry arrived, it is surprising that there has been little interest in the overall vegetation for the last 30 years, considering the expansion of ecological research. In addition to being a key milestone in the development of methodology for larger-scale monitoring programmes such as the British Countryside Survey (Carey et al., 2008) and the Welsh Glastir Monitoring and Evaluation Programme (Emmett and GMEP team, 2014), the 1974 Shetland Survey data provide a unique baseline from which changes in the vegetation could be determined from repeated locations. A repeat survey or monitoring programme based on the framework described in the present paper would not only yield important scientific results but, when interpreted using the modern procedures developed in Countryside Survey, could also be converted into policy-relevant conclusions and would add considerable value to the existing data sets. The methodology has been shown to be robust and has been used in the CS to follow changes over 30 years. CS has had considerable impact in shaping government policy in biodiversity, particularly in the realm of hedgerow legislation (Barr and Gillespie, 2000; Petit et al., 2003; Anonymous, 1997). None of this would have been achieved had the methodology not been grounded on a sound statistical base.

Within the CS project, changes taking place in the character of the vegetation can be detected, as well as their underlying causal factors. In Shetland, there are a range of factors which could have had an impact on the vegetation in the last 40 years. These include fluctuations in sheep numbers, which affect grazing intensity, and therefore vegetation composition; a decline in traditional crofting, with a possible associated loss in key habitats and species; heathland improvement schemes; peat cutting, which can damage the surface of blanket bog at intense levels; wind farm construction, which can potentially alter an area's microclimate and therefore vegetation; and also sulfur deposition from local shipping and oil deposition, particularly that spilled from the *Braer* oil tanker, which ran aground off Shetland in 1993.

In the event of a repeat survey, it would be possible to investigate any potential effects of climate change. Since the 1970s, there appears to be a trend towards higher average temperatures and wetter winters in Shetland. Monitoring the sensitive arctic–alpine species found at unusually low elevations in Shetland could serve as an early warning of the effects of climate change in a global context. The analysis of the wide range of plant types surveyed could also potentially provide information on the effects of any climate changes.

A key benefit of the repeatable methodology described is that it gives an unbiased assessment of change, as shown in the case of the CS and Woodland Survey. A repeat survey of Shetland would provide objective information about the extent of the changes in vegetation since 1974. Results from the repeat surveys and analysis of vegetation changes from the Countryside Survey (Carey et al., 2008) and the British Woodland Survey (Wood et al., 2015; Kirby et al., 2005) have both shown unexpected results that would not have been revealed, had these objective methods not been used.

Data availability

The data sets have been assigned digital object identifiers and users of the data must reference the data as follows:

Bunce, R. G. H., Bassett, P. A., and Wood, C. M.: Terrestrial habitat, vegetation and soil data from Shetland, 1974, NERC Environmental Information Data Centre, doi:10.5285/06fc0b8c-cc4a-4ea8-b4be-f8bd7ee25342, 2015.

Bunce, R. G. H. and Bassett, P. A.: Land Classification of Shetland 1974, NERC Environmental Information Data Centre, doi:10.5285/f1b3179e-b446-473d-a5fb-4166668da146, 2015.

Both of the data sets are available from the CEH Environmental Information Data Centre Gateway (https://gateway.ceh.ac.uk) and via the following links: https://catalogue.ceh.ac.uk/documents/f1b3179e-b446-473d-a5fb-4166668da146, https://catalogue.ceh.ac.uk/documents/06fc0b8c-cc4a-4ea8-b4be-f8bd7ee25342.

Data sets are provided under the terms of the Open Government Licence (http://eidchub.ceh.ac.uk/administration-folder/tools/ceh-standard-licence-texts/ceh-open-government-licence/plain, http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/)

The metadata are stored in the ISO 19115 (2003) schema (International Organization for Standardization, 2015) in the UK Gemini 2.1 profile (UK GEMINI, 2015). Users of the data sets will find the following documents useful: *Shetland Vegetation Survey Handbook of Field Methods* (Bunce, 1974) and "Report to NCC on some aspects of the ecology of Shetland. Part III: The Terrestrial Survey of Shetland" (Bunce, 1975).

Author contributions. C. M. Wood prepared the manuscript with significant contributions from R. G. H. Bunce, and is the current database manager for the Land Use Research Group at CEH Lancaster. R. G. H. Bunce designed the experiment (along with M. W. Shaw) and ran the survey in 1974.

Acknowledgements. We thank the land owners and agents who gave permission to conduct surveys on their land. Without their cooperation this project would not have been possible. We also thank the field surveyors: R. G. H. Bunce, P. Bassett, S. K. Morrell, J. Heap, J. B. Hambrey, J. Caird, J. Mackintosh, P. Bibby, M. Ferguson, M. Sturges, D. Thomas, J. Thomas, J. Wilson, P. Swain, M. Horner. Thanks also go to P. Harvey from the Shetland Amenity Trust for providing current background information regarding Shetland, to C. Hallam for entering data, and to F. Huettman and the two anonymous referees, whose comments have improved the manuscript considerably.

Edited by: F. Huettmann

References

Allot, G.: A report on the lowland vegetation of Foula, Brathay Exploration Group Field Studies Report, 11, 48–57, 1971.

Anonymous: The Hedgerows Regulations, The Stationery Office, London, 16 pp., 1997.

Barkham, J. P.: A report on the upland vegetation of Foula, Brathay Exploration Group Field Studies Report, 11, 25–47, 1971.

Barr, C. J. and Gillespie, M. K.: Estimating hedgerow length and pattern characteristics in Great Britain using Countryside Survey data, J. Environ. Manage., 60, 23–32, doi:10.1006/jema.2000.0359, 2000.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 1990 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, doi:10.5285/32dda22c-24cc-4fa8-adcb-5481b1b77ef1, 2014a.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Countryside Survey 1998 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, doi:10.5285/ad7babc3-6b43-4754-8981-edcf03769f11, 2014b.

Benefield, C. B. and Bunce, R. G. H.: A preliminary visual presentation of land classes in Britain, Institute of Terrestrial Ecology, Grange over Sands, 39 pp., 1982.

Berry, R. J. and Johnston, J. L.: The natural history of Shetland, Collins, London, 1980.

Birse, E. L.: Bioclimatic characteristics of Shetland, in: The Natural Environment of Shetland, edited by: Goodier, R., NCC, Edinburgh, 24–32, 1974.

Blockeel, T. L., Bosanquet, S. D., Hill, M., and Preston, C. D.: Atlas of British & Irish Bryophytes, Pisces Publications, Newbury, 2014.

British Geological Survey: Quarter-inch geological map of Scotland HMSO, Director General of the Ordnance Survey, Chessington, Surrey, 1963.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 2007 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, doi:10.5285/f03cba75-8bca-4679-ae2a-77a9fcbd4df3, 2014.

Bunce, R. G. H.: Shetland Vegetation Survey Handbook of Field Methods, Institute of Terrestrial Ecology, Grange over Sands, 28 pp., 1974.

Bunce, R. G. H.: Report to NCC on some aspects of the ecology of Shetland. Part III: The Terrestrial Survey of Shetland, Institute of Terrestrial Ecology, Grange over Sands, 1975.

Bunce, R. G. H. and Shaw, M. W.: A standardised method for ecological survey, J. Environ. Manage., 1, 239–258, 1973.

Bunce, R. G. H. and Smith, R. S.: An ecological survey of Cumbria, Structure Plan Working Paper, 4, Cumbria County Council and Lake District Special Planning Board, Kendal, 1978.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE land classification of Great Britain 1990, NERC

Environmental Information Data Centre, doi:10.5285/ab320e08-faf5-48e1-9ec9-77a213d2907f, 1990.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE Merlewood Land Classification of Great Britain, J. Biogeogr., 23, 625–634, 1996a.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: Land classification for strategic ecological survey, J. Environ. Manage., 47, 37–60, doi:10.1006/jema.1996.0034, 1996b.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D., and Lane, A. M. J.: ITE Land Classification of Great Britain 1998, NERC Environmental Information Data Centre, doi:10.5285/971671a6-98b4-4d80-b165-21dace7373b9, 1998.

Bunce, R. G. H., Groom, G. B., Jongman, R. H. G., Padoa-Schioppa, E., Howard, D. C., and Petit, S.: Handbook for surveillance and monitoring of European Habitats. First edition, edited by: Bunce, R. G. H., Groom, G. B., Jongman, R. H. G., and Padoa-Schioppa, E., Wageningen, Alterra (Alterra report 1219), 107 pp., 2005.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D., and Scott, A.: ITE Land Classification of Great Britain 2007, NERC Environmental Information Data Centre, doi:10.5285/5f0605e4-aa2a-48ab-b47c-bf5510823e8f, 2007.

Bunce, R. G. H., Metzger, M. J., Jongman, R. H. G., Brandt, J., de Blust, G., Elena-Rossello, R., Groom, G. B., Halada, L., Hofer, G., Howard, D. C., Kovář, P., Mücher, C. A., Padoa-Schioppa, E., Paelinx, D., Palo, A., Perez-Soba, M., Ramos, I. L., Roche, P., Skånes, H., and Wrbka, T.: A standardized procedure for surveillance and monitoring European habitats and provision of spatial data, Landsc. Ecol., 23, 11–25, doi:10.1007/s10980-007-9173-8, 2008.

Bunce, R. G. H., Bogers, M., and Roche, P.: Manual for habitat and vegetation surveillance and monitoring: temperate, mediterranean and desert biomes, Alterra, Wageningen, 2011.

Bunce, R. G. H., Wood, C. M., Henrys, P. A., Smart, S. M., Howard, D. C., and Barr, C. J.: Countryside Survey 1978 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, doi:10.5285/c8feabac-7c3f-4c35-91f7-296782b957d1, 2012.

Bunce, R. G. H. and Bassett, P. A.: Land Classification of Shetland 1974, NERC Environmental Information Data Centre, doi:10.5285/f1b3179e-b446-473d-a5fb-4166668da146, 2015.

Bunce, R. G. H., Bassett, P. A., and Wood, C. M.: Terrestrial habitat, vegetation and soil data from Shetland, 1974, NERC Environmental Information Data Centre, doi:10.5285/06fc0b8c-cc4a-4ea8-b4be-f8bd7ee25342, 2015.

Carey, P. D., Wallis, S., Chamberlain, P. M., Cooper, A., Emmett, B. A., Maskell, L. C., McCann, T., Murphy, J., Norton, L. R., Reynolds, B., Scott, W. A., Simpson, I. C., Smart, S. M., and Ullyett, J. M.: Countryside Survey: UK Results from 2007, NERC/Centre for Ecology & Hydrology, Lancaster, 2008.

Chambers, N., Simmons, C., and Wackernagel, M.: Sharing nature's interest: ecological footprints as an indicator of sustainability, Routledge, Oxford & New York, 2014.

Clapham, A. R., Tutin, T. G., and Warburg, E. F.: Flora of the British Isles, Cambridge University Press, Cambridge, 1952.

Cranfield University: National Soils Inventory, available at: http://www.landis.org.uk/about/index.cfm, last access: 2 September 2015.

Czech, B.: Technological progress and biodiversity conservation: a dollar spent, a dollar burned, Conserv. Biol., 17, 1455–1457, 2003.

Dalby, D. H. and Dalby, C.: Shetland Lichens, Shetland Amenity Trust, Lerwick, 2005.

Daly, H. E. and Farley, J.: Ecological economics: principles and applications, Island Press, 2011.

Emmett, B. E. and GMEP team: Glastir Monitoring & Evaluation Programme, First Year Annual Report to Welsh Government (Contract reference: C147/2010/11), NERC/Centre for Ecology & Hydrology, Bangor, 442, 2014.

Forestry Commission: National Forest Inventory, available at: http://www.forestry.gov.uk/website/forestry.nsf/byunique/infd-89s9ls, last access: 15 January 2013.

Goode, D.: The flora and vegetation of Shetland, in: The Natural Environment of Shetland, edited by: Goodier, R., NCC, Edinburgh, 50–72, 1974.

Grabherr, G., Gottfried, M., and Pauli, H.: Climate change impacts in alpine environments, Geography Compass, 4, 1133–1153, doi:10.1111/j.1749-8198.2010.00356.x, 2010.

Hill, M. and Paton, J.: The Bryophytes of Shetland., Nature Conservancy Council/Institute of Terrestrial Ecology, Bangor, 1976.

Hill, M. O.: Reciprocal averaging: an eigenvector method of ordination, The J. Ecol., 61, 237–249, doi:10.2307/2258931, 1973.

Hill, M. O. and Šmilauer, P.: TWINSPAN for Windows version 2.3, Centre for Ecology and Hydrology, University of South Bohemia, Wallingford, 2005.

Hill, M. O., Bunce, R. G. H., and Shaw, M. W.: Indicator species analysis, a divisive polythetic method of classification, and its application to a survey of native pinewoods in Scotland, J. Ecol., 63, 597–613, doi:10.2307/2258738, 1975.

Hill, M. O., Preston, C. D., and Smith, A. J. E.: Atlas of the bryophytes of Britain and Ireland, Harley books, Colchester, 1991.

Huettmann, F.: An Oceanography Conservation View of Central America: Caribbean Wealth, Pacific Wilderness, Plunder and Mismanagement Meets Seafloor Mining, Deep-Sea Drilling, Climate Change, and Human Population Explosion in the EEZs and Beyond, in: Central American Biodiversity: Conservation, Ecology, and a Sustainable Future, edited by: Huettmann, F., Springer New York, New York, NY, 135-158, 2015.

International Organization for Standardization: ISO 19115 (2003), available at: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020, last access: 12 October 2015.

Jackson, D.: Guidance on the Interpretation of the Biodiversity Broad Habitat Classification (Terrestrial and Freshwater Types): Definitions and the Relationship with Other Habitat Classifications (JNCC Report, No. 307), JNCC, 73 pp., 2000.

Johnston, J. L.: A naturalist's Shetland, T. and A.D. Poyser, London, 1999.

Kirby, K. J., Smart, S. M., Black, H. I. J., Bunce, R. G. H., Corney, P. M., and Smithers, R. J.: Long term ecological change in British woodland (1971–2001), A re-survey and analysis of change based on the 103 sites in the Nature Conservancy "Bunce 1971" woodland survey, Final report, Peterborough: English Nature (English Nature Research Reports Number 653), 139 + appendices, 2005.

Lautenbacher, C. C.: The Global Earth Observation System of Systems: Science Serving Society, Space Policy, 22, 8–11, doi:10.1016/j.spacepol.2005.12.004, 2006.

Lengyel, S., Déri, E., Varga, Z., Horváth, R., Tóthmérész, B., Henry, P.-Y., Kobler, A., Kutnar, L., Babij, V., and Seliškar, A.: Habitat monitoring in Europe: a description of current practices, Biodivers. Conserv., 17, 3327–3339, doi:10.1007/s10531-008-9395-3, 2008a.

Lengyel, S., Kobler, A., Kutnar, L., Framstad, E., Henry, P.-Y., Babij, V., Gruber, B., Schmeller, D., and Henle, K.: A review and a framework for the integration of biodiversity monitoring at the habitat level, Biodivers. Conserv., 17, 3341–3356, doi:10.1007/s10531-008-9359-7, 2008b.

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., and Barr, C. J.: Vegetation Plots Handbook CS Technical Report No.2/07, Centre for Ecology and Hydrology, Lancaster, 2008a.

Maskell, L. C., Norton, L. R., Smart, S. M., Carey, P. D., Murphy, J., Chamberlain, P. M., Wood, C. M., Bunce, R. G. H., and Barr, C. J.: Countryside Survey. Field Mapping Handbook CS Technical Report No.1/07, Centre for Ecology and Hydrology, Lancaster, 2008b.

McVean, D. N. and Ratcliffe, D. A.: Plant communities of the Scottish Highlands, A study of Scottish mountain, moorland and forest vegetation, Monographs of the Nature Conservancy, HMSO, London, 445 pp., 1962.

Met Office: UK climate – Historic station data (Lerwick), available at: http://www.metoffice.gov.uk/pub/data/weather/uk/climate/stationdata/lerwickdata.txt, last access: 1 September 2015.

Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Mücher, C. A., and Watkins, J. W.: A climatic stratification of the environment of Europe, Global Ecol. Biogeogr., 14, 549–563, doi:10.1111/j.1466-822X.2005.00190.x, 2005.

Metzger, M., Bunce, R., Leemans, R., and Viner, D.: Projected environmental shifts under climate change: European trends and regional impacts, Environ. Conserv., 35, 64–75, doi:10.1017/S0376892908004529, 2008.

Metzger, M. J., Brus, D. J., Bunce, R. G. H., Carey, P. D., Gonçalves, J., Honrado, J. P., Jongman, R. H. G., Trabucco, A., and Zomer, R.: Environmental stratifications as the basis for national, European and global ecological monitoring, Ecol. Indic., 33, 26–35, doi:10.1016/j.ecolind.2012.11.009, 2013a.

Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Sayre, R., Trabucco, A., and Zomer, R.: A high-resolution bioclimate map of the world: a unifying framework for global biodiversity research and monitoring, Global Ecol. Biogeogr., 22, 630–638, doi:10.1111/geb.12022, 2013b.

Milner, C.: Shetland project monitoring report Nature Conservancy Council, Contract Report to the Nature Conservancy Council, 1975.

Norton, L. R., Maskell, L. C., Smart, S. S., Dunbar, M. J., Emmett, B. A., Carey, P. D., Williams, P., Crowe, A., Chandler, K., Scott, W. A., and Wood, C. M.: Measuring stock and change in the GB countryside for policy: key findings and developments from the Countryside Survey 2007 field survey, J. Environ. Manage., 113, 117–127, doi:10.1016/j.jenvman.2012.07.030, 2012.

Ott, R.: Sound truth and corporate myth$: the legacy of the Exxon Valdez oil spill, Dragonfly Sisters Press, Cordova, 2005.

Parr, T., Ferretti, M., Simpson, I., Forsius, M., and Kovacs-Lang, E.: Towards a long-term integrated monitoring programme in Europe: network design in theory and practice, Environ. Monit. Assess., 78, 253–290, 2002.

Parr, T. W., Sier, A. R. J., Battarbee, R. W., Mackay, A., and Burgess, J.: Detecting environmental change: science and society – perspectives on long-term research and monitoring in the 21st century, Sci. Total Environ., 310, 1–8, doi:10.1016/S0048-9697(03)00257-2, 2003.

Petit, S., Stuart, R., Gillespie, M., and Barr, C.: Field boundaries in Great Britain: stock and change between 1984, 1990 and 1998, J. Environ. Manage., 67, 229–238, doi:10.1016/S0301-4797(02)00176-7, 2003.

Preston, C. D., Pearman, D. A., and Dines, T. D.: New atlas of the British and Irish flora, An atlas of the vascular plants of Britain, Ireland, the Isle of Man and the Channel Islands, Oxford University Press, Oxford, 2002.

Roche, P. and Geijzendorffer, I.: EBONE: integrated figures of habitat and biodiversity indicators: quantifying indicators of an integrated biodiversity observation system, Alterra, Wageningen UR, 2013.

Roper-Lindsay, J. and Say, A.: Plant communities of the Shetland Islands, J. Ecol., 74, 1013–1030, doi:10.2307/2260230, 1986.

Scott, W. and Palmer, R.: The flowering plants and ferns of the Shetland Islands, Lerwick: Shetland Times Ltd ix, 468p.-col. illus., maps., ISBN 900662565, 1987.

Scott, W. A.: Countryside Survey. Statistical Report (Countryside Survey Technical Report No. 4/07), NERC Centre for Ecology and Hydrology Lancaster, 2008.

Scottish Government: Scottish Agricultural Census, available at: http://www.gov.scot/Topics/Statistics/Browse/Agriculture-Fisheries/PubScottishCensus, last access: 7 January 2016.

Scottish Natural Heritage: Shetland, Scottish Natural Heritage, Perth, 2002.

Scottish Rural Development Programme: What Can SRDP Do For Crofting And Small Units?, available at: http://www.crofting.scotland.gov.uk/userfiles/documents/SRDP%20paper%20final.pdf, last access: 7 January 2016.

Sheail, J. and Bunce, R. G. H.: The development and scientific principles of an environmental classification for strategic ecological survey in the United Kingdom, Environ. Conserv., 30, 147–159, doi:10.1017/S0376892903000134, 2003.

Smart, S., Maskell, L., Dunbar, M., Emmett, B., Marks, S., Norton, L., Rose, P., and Simpson, I.: An Integrated Assessment of Countryside Survey data to investigate Ecosystem Services in Great Britain, Countryside Survey Technical Report No. 10/07, NERC Centre for Ecology and Hydrology, Lancaster, 2010.

Spence, D. H. N.: Subarctic debris and scrub vegetation in Shetland, in: The Natural Environment of Shetland, edited by: Goodier, R., NCC, Edinburgh, 50–72, 1974.

UK Environmental Change Network: UK Environmental Change Network, available at: http://www.ecn.ac.uk/, last access: 2 September 2015.

UK GEMINI: UK GEMINI, available at: http://www.agi.org.uk/join-us/agi-groups/standards-committee/uk-gemini, last access: 12 May 2015.

Viking Energy: The Project (The Viking Wind Farm), available at: http://www.vikingenergy.co.uk/the-project, last access: 7 January 2016.

Wills, J.: A place in the Sun: Shetland and Oil-Myths and Realities, Mainstream Publishing Company, Edinburgh, 1991.

Wilson, E.: Freedom and loss in a human landscape: Multinational oil exploitation and the survival of reindeer herding in north-eastern Sakhalin, the Russian Far East, Sibirica, 3, 21–48, doi:10.1080/1361736032000168012, 2003.

Wood, C. M., Smart, S. M., and Bunce, R. G. H.: Woodland Survey of Great Britain 1971–2001, Earth Syst. Sci. Data, 7, 203–214, doi:10.5194/essd-7-203-2015, 2015.

Earth Syst. Sci. Data, 10, 745–763, 2018
https://doi.org/10.5194/essd-10-745-2018

Open Access Earth System Science Data

Ecological landscape elements: long-term monitoring in Great Britain, the Countryside Survey 1978–2007 and beyond

Claire M. Wood[1], **Robert G. H. Bunce**[2], **Lisa R. Norton**[1], **Lindsay C. Maskell**[1], **Simon M. Smart**[1], **W. Andrew Scott**[1], **Peter A. Henrys**[1], **David C. Howard**[1], **Simon M. Wright**[1], **Michael J. Brown**[1], **Rod J. Scott**[1], **Rick C. Stuart**[1], and **John W. Watkins**[1]

[1]Centre for Ecology & Hydrology, Lancaster Environment Centre, Bailrigg, Lancaster, LA1 4AP, UK
[2]Estonian University of Life Sciences, Kreuzwaldi 5, 51014 Tartu, Estonia

Correspondence: Claire M. Wood (clamw@ceh.ac.uk)

Received: 25 October 2017 – Discussion started: 8 November 2017
Revised: 22 March 2018 – Accepted: 22 March 2018 – Published: 16 April 2018

Abstract. The Countryside Survey (CS) of Great Britain (GB) provides a unique and statistically robust series of datasets, consisting of an extensive set of repeated ecological measurements at a national scale, covering a time span of 29 years. CS was first undertaken in 1978 to provide a baseline for ecological and land use change monitoring in the rural environment of GB, following a stratified random design, based on 1 km squares. Originally, eight random 1 km squares were drawn from each of 32 environmental classes, thus comprising 256 sample squares in the 1978 survey. The number of these sites increased to 382 in 1984, 506 in 1990, 569 in 1998 and 591 in 2007. Detailed information regarding vegetation types and land use was mapped in all five surveys, allowing reporting by defined standard habitat classifications. Additionally, point and linear landscape features (such as trees and hedgerows) are available from all surveys after 1978. From these stratified, randomly located sample squares, information can be converted into national estimates, with associated error terms.

Other data, relating to soils, freshwater and vegetation, were also sampled on analogous dates. However, the present paper describes only the surveys of landscape features and habitats. The resulting datasets provide a unique, comprehensive, quantitative ecological coverage of extent and change in these features in GB. Basic results are presented and their implications discussed. However, much opportunity for further analyses remains.

Data from each of the survey years are available via the following DOIs: Landscape area data 1978: https://doi.org/10.5285/86c017ba-dc62-46f0-ad13-c862bf31740e, 1984: https://doi.org/10.5285/b656bb43-448d-4b2c-aade-7993aa243ea3, 1990: https://doi.org/10.5285/94f664e5-10f2-4655-bfe6-44d745f5dca7, 1998: https://doi.org/10.5285/1e050028-5c55-42f4-a0ea-c895d827b824, and 2007: https://doi.org/10.5285/bf189c57-61eb-4339-a7b3-d2e81fdde28d; Landscape linear feature data 1984: https://doi.org/10.5285/a3f5665c-94b2-4c46-909e-a98be97857e5, 1990: https://doi.org/10.5285/311daad4-bc8c-485a-bc8a-e0d054889219, 1998: https://doi.org/10.5285/8aaf6f8c-c245-46bb-8a2a-f0db012b2643 and 2007: https://doi.org/10.5285/e1d31245-4c0a-4dee-b36c-b23f1a697f88, Landscape point feature data 1984: https://doi.org/10.5285/124b872e-036e-4dd3-8316-476b5f42c16e, 1990: https://doi.org/10.5285/1481bc63-80d7-4d18-bcba-8804aa0a9e1b, 1998: https://doi.org/10.5285/ed10944f-40c8-4913-b3f5-13c8e844e153 and 2007: https://doi.org/10.5285/55dc5fd7-d3f7-4440-b8a7-7187f8b0550b.

Published by Copernicus Publications.

1 Introduction

The Countryside Survey (CS) of Great Britain (GB) was initiated in the late 1970s for the surveillance and monitoring of ecological and land cover change in the rural environment using quantitative and repeatable methods. Retaining standardised methods to describe the habitats, landscape features, land use, soils, freshwater and vegetation present, has allowed data for subsequent surveys to estimate change. The survey provides a wealth of ecological data, consisting of a detailed range of measurements at a national scale, covering five surveys across a time span of 29 years (1978–2007), with the intention of future repeat surveys. The history of the development of the methodology is given by Sheail and Bunce (2003). A number of simultaneous surveys have also been undertaken in Northern Ireland (Cooper et al., 2009), complementing the GB survey, and enabling reporting for the United Kingdom (UK) as whole. However, the data from GB are the focus of this paper.

The survey is based on 1 km squares as a conveniently sized unit for landscape monitoring. This had previously been tested in Cumbria (1975) (Bunce and Smith, 1978) and Shetland (1974) (Wood and Bunce, 2016) in the years preceding the first GB survey in 1978. The survey design is based on a series of distributed, stratified, randomly selected 1 km sample squares from across Britain, which numbered 256 in 1978, 382 in 1984, 506 in 1990, 569 in 1998 and 591 in 2007 (Fig. 1). The stratification used is the ITE Land Classification of 1 km squares in GB (Fig. 2), which is based on a statistical analysis of topographic, physiographic and climatic attributes as described in (Bunce et al., 1996a, c) and summarised in Sect. 2.

The most geographically comprehensive element of the survey is the mapping of land cover and ecologically relevant landscape features, carried out in every survey undertaken thus far (1978, 1984, 1990, 1998, and most recently, 2007). Across survey areas of 1 km square, area, line and point features are mapped onto base maps, using a range of pre-determined coded options. Areas are categorised by predominant vegetation characteristics and, in 2007, were assigned to broad and priority habitats (Jackson, 2000; Maddock, 2008), which are able to be translated into the habitats of Annex 1 of the EU Habitats directive (Romão, 2013). Mapping was initially carried out using waterproof paper base maps, but for the first time in 2007, data were collected in digital format using rugged field computers.

With the inclusion of the vegetation data (as described in Wood et al., 2017), soils and freshwater data (Emmett et al., 2010; Dunbar et al., 2010; Williams et al., 2010; Carey et al., 2008), the survey as a whole provides a wide range of nationally significant ecological datasets, globally unique in their geographical coverage and time span. The co-registration of all the data, in both time and space, along with the flexibility in coding make the datasets unique in describing and interpreting the drivers of change in the British landscape. In particular, other examples of field-mapped land cover data sets, with their potential for assessing detailed changes in countryside structure at a national level, are not known to the authors. The majority of other field habitat mapping projects are one-off exercises which are not intended to monitor change and do not use repeatable methods. Monitoring requires more stringent procedures to ensure that differences recorded represent real change and not distortions due to differences between observers or recording technique, as described by Brandt et al. (2002). One commonly used option for landscape mapping is the use of large-scale land cover maps, largely derived from satellite or aerial imagery (Cole et al., 2015; Mayaux et al., 2004; Eva et al., 2004; Bartholomé and Belward, 2005). None of these examples include the same level of detail, with the same potential for assessing change or integrating with co-located in situ data, over such a time span as the data from the Countryside Survey. Whilst the CS field data are complemented by a series of land cover maps (Morton et al., 2011; Fuller et al., 1994a, 2001), which are useful for determining habitat extent, they do not provide data to determine habitat quality and condition, habitat change or the extent and condition of landscape point and line features.

2 Survey design: site selection and stratification

Following preliminary work undertaken within smaller regions of Britain (Wood and Bunce, 2016; Bunce and Smith, 1978), a sample unit of 1 km square was found to be an effective size for capturing data within CS. A 1 km square is small enough to survey in a relatively short period of time (1 week or less) and yet large enough to contain sufficient environmental features to allow differentiation of the character of squares, and interaction between components to be examined.

With over 240 000 one-kilometre squares in GB, a sampling approach was essential and a statistical environmental classification was constructed from which stratified, random samples were taken. This classification covered the whole of Great Britain using multivariate analysis of environmental factors, for example altitude and climate (converted into attributes which the statistical methods at the time could analyse) from each 1 km square (Bunce et al., 1996b). A primary objective of this methodology was to minimise bias, as the classification divides the population into discrete strata that are then used to derive samples from which ecological parameters such as vegetation can be recorded. By using this statistically robust method, it is then possible to scale up the results from the sample sites to describe the entire population, with associated error terms.

The sampling methodology was initially developed at regional scales in the early 1970s, for example in Shetland and Cumbria (Bunce and Smith, 1978; Wood and Bunce, 2016). Later it was extended to the whole of GB, but only

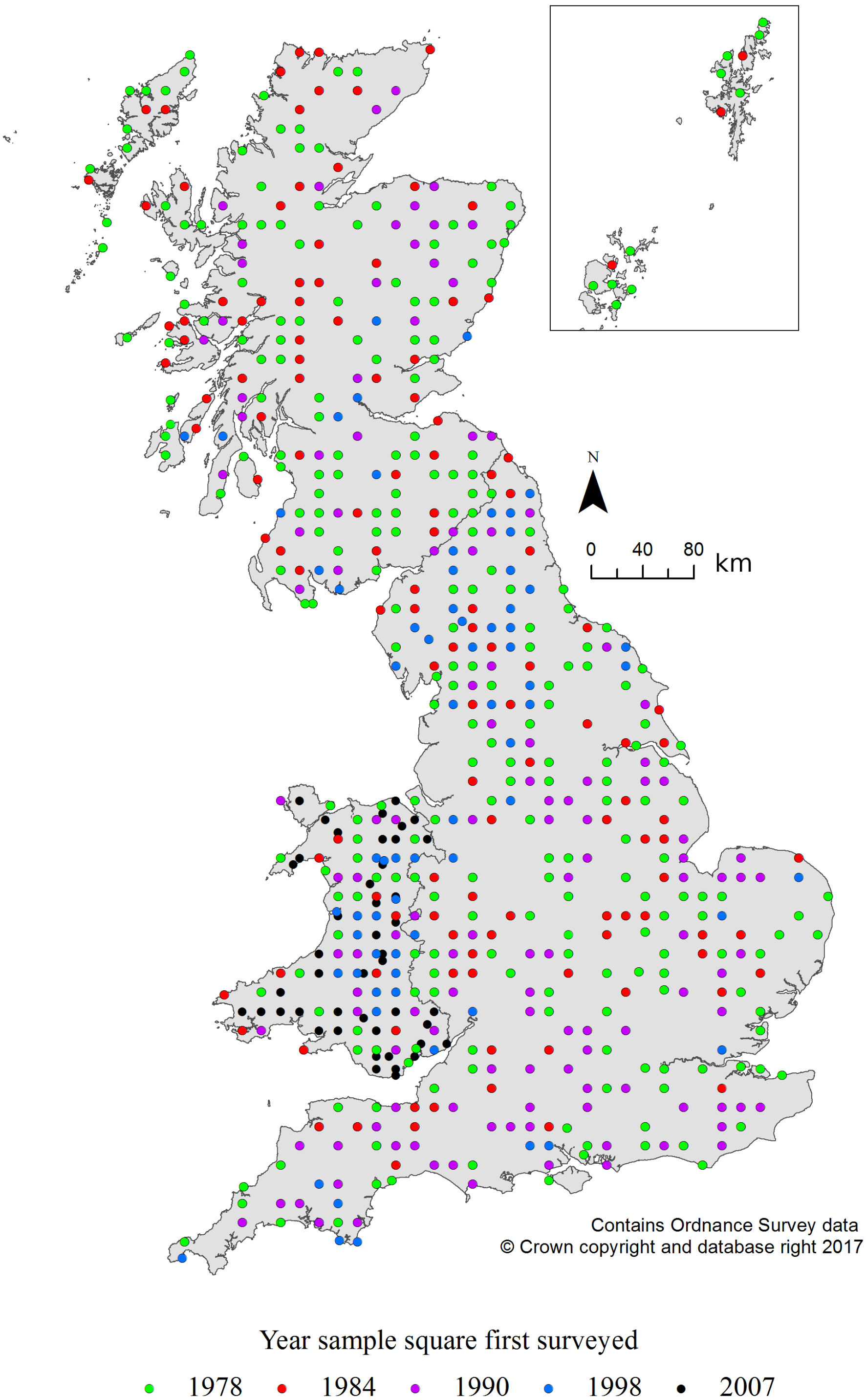


Figure 1. Map of sampling locations across Britain.

on a grid of 1225 × 1 km squares as a consequence of the limitations of computing power at the time. By the end of the 1980s, all 1 km squares in GB had been classified into the same 32 strata, which was not technically possible at the start of the 1970s. Known as the "Institute of Terrestrial Ecology (ITE) Land Classification of Great Britain" (Bunce et al., 1990, 1996a, b), it has evolved over the 30-year period (Sheail and Bunce, 2003) latterly to allow the reporting of separate national estimates for Scotland (1998), and then Wales (2007). However, changes in the stratification have all been conservative so as not to compromise previous work, and the basic stratification still underpins CS, now with

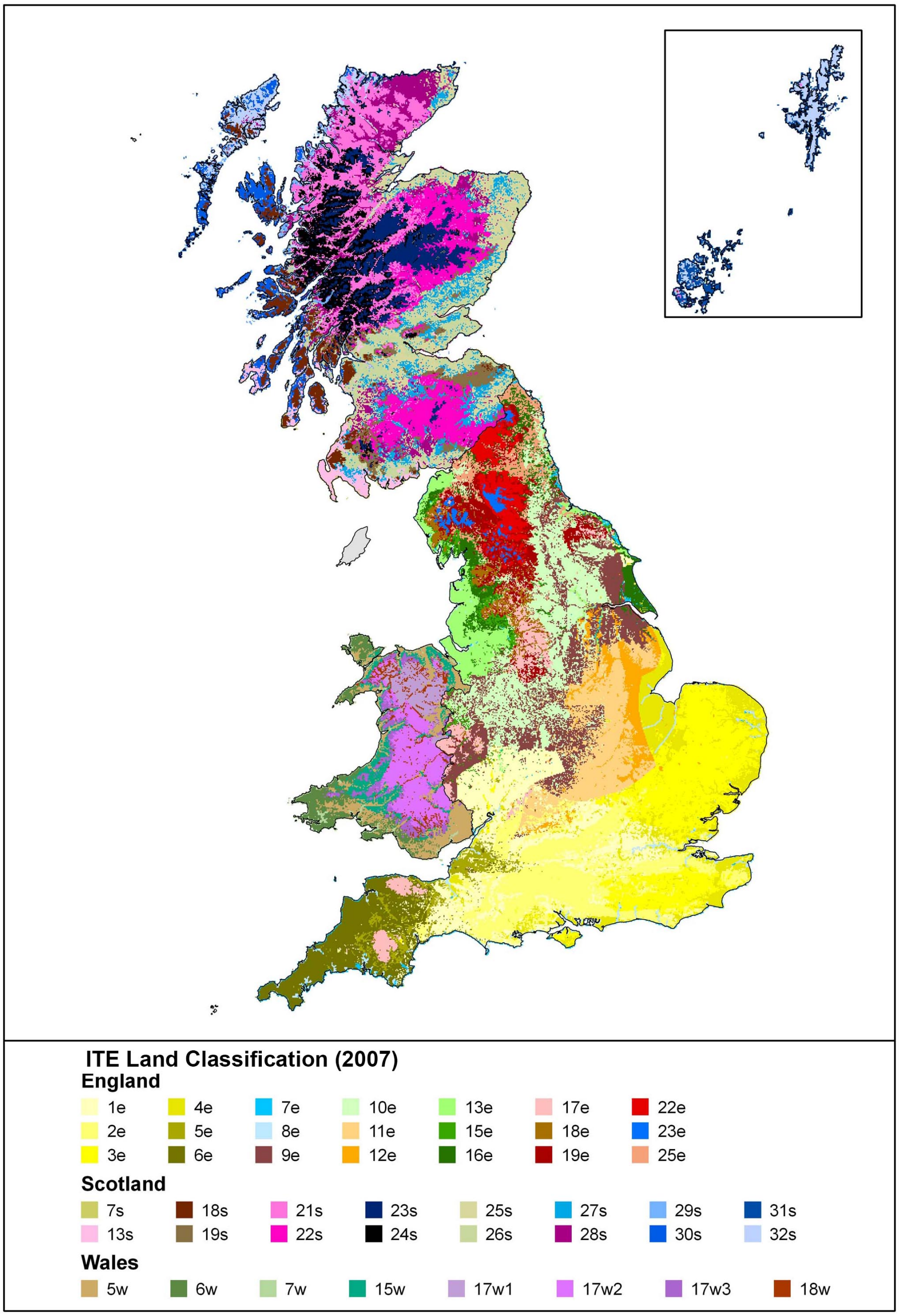


Figure 2. ITE Land Classification, 2007. (Contains Ordnance Survey data (©) Crown copyright and database right 2018.)

45 strata (or "land classes"). Further details are provided in Wood et al. (2017) and Barr and Wood (2011). This successful method of the consistent classification of land into relatively homogenous strata has been proven to provide a valuable spatial framework as the basis for monitoring ecological indicators across large areas. There are now several examples of where the British methods have been emulated effectively, including Northern Ireland (Cooper, 2000), Spain (Elena-Rosselló, 1997), Norway (Bakkestuen et al., 2008), Sweden (Ståhl et al., 2011), Estonia (Villoslada et al., 2016) and Europe (and the whole world) (Metzger et al., 2013).

2.1 Sampling sites

Having generated the classification to act as the sampling stratification system, the number of samples to be surveyed in the first (1978) survey was considered. Ideally, this number would depend on the size of the stratum (i.e. how many 1 km

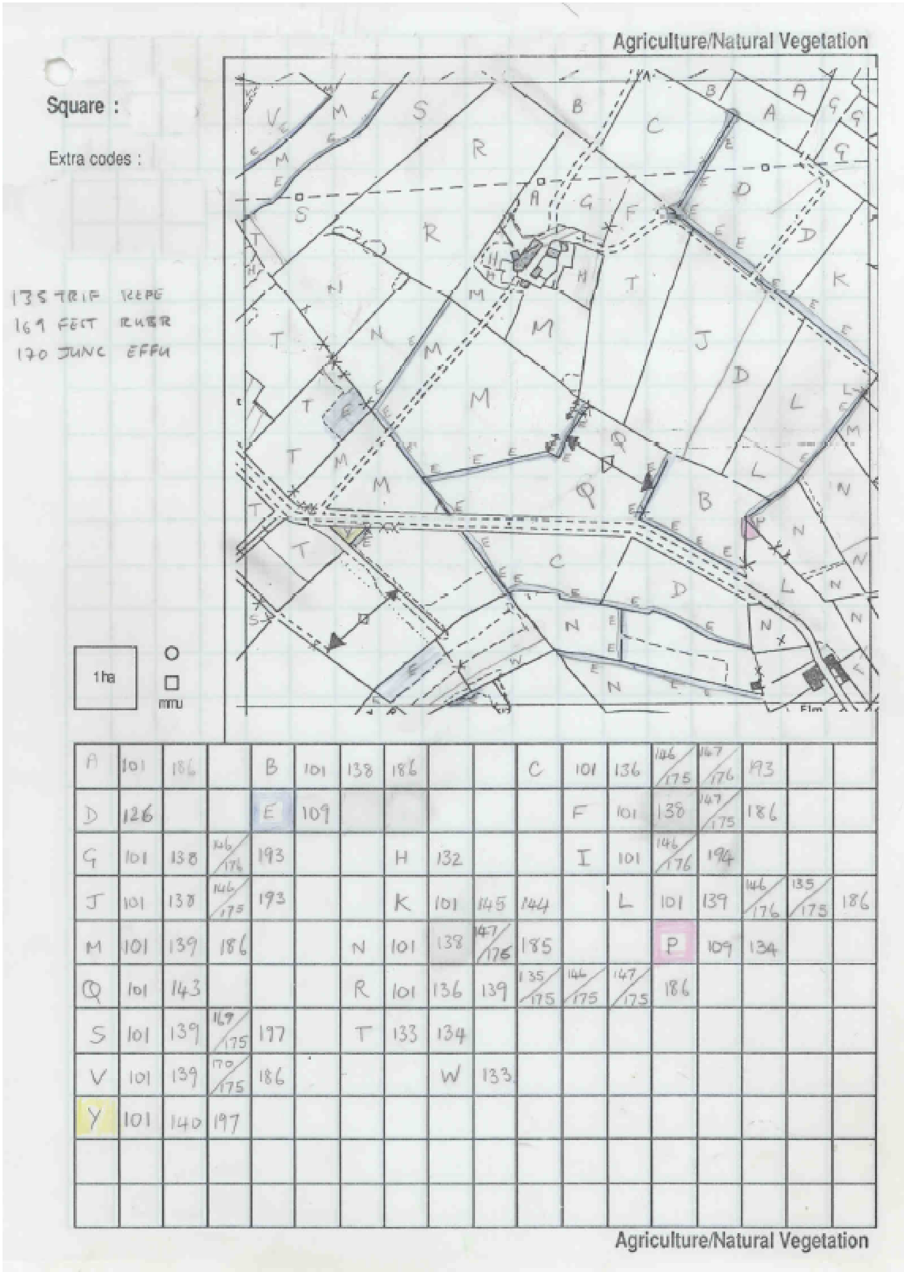


Figure 3. Example of a field mapping sheet.

squares of the land class occurred in GB) and on the ecological variability within the stratum. Preliminary work had suggested that, for ecological surveys of this type, at least eight samples per stratum were necessary in order to be representative of that stratum. As resources were constrained, eight squares were thus selected at random from each of the strata/land classes. These squares were taken from the grid of classified squares and thus the final sample for the first GB survey was a gridded, stratified, random sample of 256 one-kilometre squares. The survey was carried out in the summers of 1977 (when a few pilot squares were sampled) and 1978 and focused on vegetation quadrats and soils; habitat areas were also mapped. In subsequent surveys, the number of 1 km squares sampled increased with each survey to a total of 591 in 2007. By 2007, the number of squares sampled from within each land class varied with the size of the stratum. The majority of squares have been repeated in each survey. However, occasionally a square may have been refused access by the landowner, in which case a replacement square would be chosen at random to replace the refused square, from within the same land class. The key requirement is that enough squares are surveyed from within each land class in

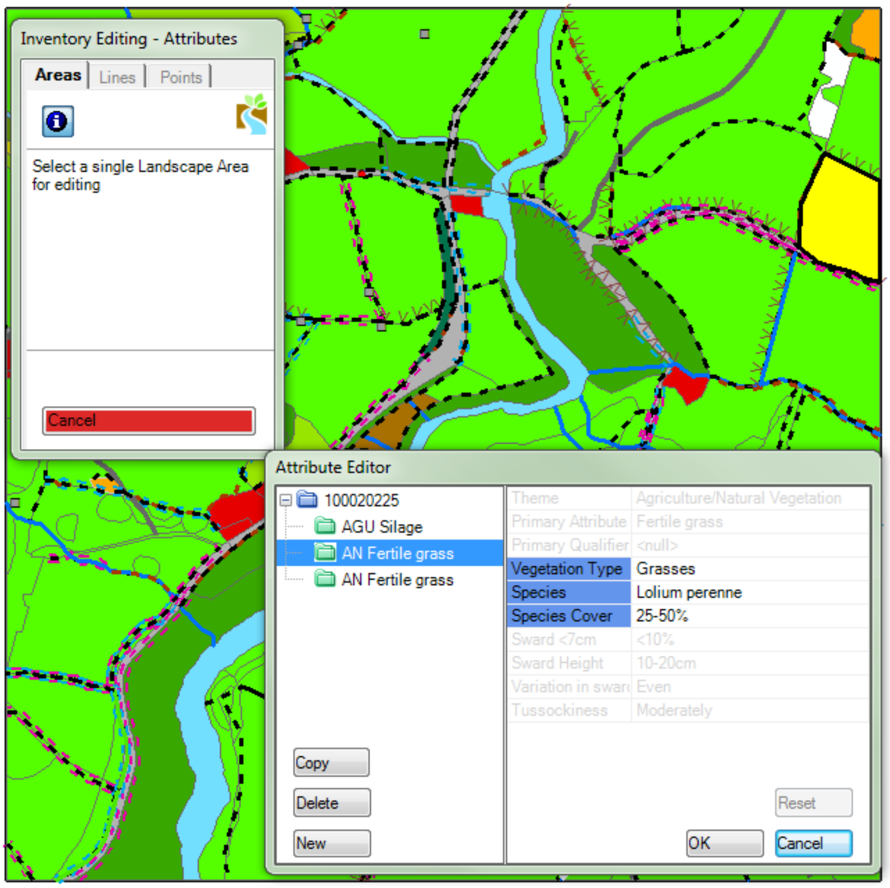


Figure 4. Example of digital mapping interface.

order to provide statistically valid estimates of features for each land class area. These issues are described in detail in Barr and Wood (2011).

2.2 Data collection methods

The mapping component of CS was carried out in the 1978, 1984, 1990, 1998 and 2007 surveys. When the methods for CS were first tested in the 1970s, the earlier regional surveys relied on vegetation plots to sample the habitats, features and vegetation types in question, together with a record of description codes, as described in Wood and Bunce (2016). By 1978, it was realised that plot sampling points alone were failing to capture the range of land cover variation within squares, and mapping the land cover across the square would address this issue. Whilst the five distributed randomly located vegetation plots (per 1 km square), with additional plots to represent rivers, roads and hedges are a key aspect of the survey for measuring habitat quality, field mapping was introduced to the survey to obtain more robust estimates of habitat extent. Predetermined standard codes, described below, were used to define land cover categories which could be converted into habitat classes. Within each survey area of 1 km square, areal, line and point features are mapped onto base maps, using a range of pre-determined coded options, using the methodologies outlined below. For the 1978, 1984, 1990 and 1998 surveys, features were mapped onto a range of paper base maps (see Fig. 3 for an example from 1990), arranged into "themes" (for example, "agriculture and natural vegetation", "forestry", "structures"). In 2007, surveyors used electronic data capture equipment for the first time in CS, and new electronic mapping software ("CS Surveyor") was developed by the Centre for Ecology & Hydrology, in conjunction with the GIS software company Esri UK. This enabled all features to be mapped onto the same, digital, base map (Fig. 4), allowed more rapid reporting of results and allowed for validation in the field.

3 Mapped features

3.1 Area features

In the first survey in 1978, areas of distinct land cover types were drawn on base maps, then later transferred onto Ordnance Survey (the UK's national mapping agency) 1 : 10 000 base maps using a set of 80 codes (see Wood et al., 2012). Cover types were mainly differentiated using dominant plant species, reflecting traditionally taught divisions between habitats based on indicator species. In the 1984 survey, the 1 : 10 000 base maps were annotated with an updated set of codes, but maintaining the integrity of the previous definitions. Parcels and features were labelled on the map with alphabetic codes, and a set of numeric feature codes were recorded against each alphabetic map code. The surveyors entered information about each mapped polygon, including land use (crop, grazing animals etc.) and at least the two most common species. The full code list is given in Barr (1984). The approach in 1990 was similar to 1984, again with a slightly updated codes list (Barr, 1990). Updating of codes reflected experience of habitat combinations gained in the field from previous surveys.

Methods of classifying land cover types in GB evolved (for example, see Wyatt et al., 1994) and immediately prior to the 1998 CS, the broad habitat system was devised and introduced by the Joint Nature Conservation Committee (Jackson, 2000). As the surveys in 1978, 1984 and in 1990 were carried out using earlier definitions for habitats that were not directly related to the broad habitat classification, a translation protocol had to be developed to ensure that past data would remain valuable for investigating change. The 1978 data were translated to broad habitats in 2009 (Wood et al., 2012). In a few cases, the translation of the 1978, 1984 and 1990 into the broad habitats has not been straightforward, due to inherent overlap between habitat classes. It had been possible, however, due to the nature of the codes recorded in the field which are recorded in a disaggregated manner, allowing the translation of parcels into different reporting categories, such as the broad habitats, or for example European Annex 1 categories (Bunce et al., 2013). In the case of the broad habitat categories, translation errors have been minimised by checks for consistency across surveys.

Broad and priority habitats were identified in the field using a key developed in 1998 (for the 1998 survey) and updated with improvements between 2001 and 2006 in time for the 2007 survey (see Maskell et al., 2008) utilising advice from many experts on UK Biodiversity Action Plan (BAP) definitions. The habitat key provides detailed guidance on how to assign areas to habitat classes using indicator species as well as physiography and dominant plant cover. It includes a key to the newly classified priority habitats (Maddock, 2008) which were incorporated in the 2007 survey; polygons assigned to a priority habitat in 2007 could be “back-allocated” to 1998 if the surveyor judged the patch had not changed across that time period. As well as mapping priority habitats, observers recorded associated species and were also encouraged to place a 2 m × 2 m targeted sampling plot in each priority habitat if it did not already have an existing plot located in it (Wood et al., 2017). For the earlier surveys, priority habitats were assigned where there was an existing habitat code that matched the current definition (for example, “Coastal Saltmarsh”).

All features were mapped using a minimum mappable unit (MMU) of 400 m^2 (20 m × 20 m through to 80 m × 5 m). No habitat was mapped as a separate unit unless it had at least this extent. If surveyors felt that an important feature was not being captured they could either create mosaics of different habitats or map elements as points or lines; clear instructions were set out the field handbooks.

In the 1998 and 2007 surveys, surveyors concentrated particularly on identifying and mapping where change had taken place between surveys, with surveyors referring to the previously mapped data when surveying (with the exception of squares being mapped for the first time).

3.2 Linear features

Linear features are landscape elements less than 5 m wide that form lines in the landscape. CS records the length and condition of a range of linear features predominantly, but not exclusively, describing boundaries. These include managed woody linear features (i.e. hedges), unmanaged woody linear features (i.e. lines of trees), walls, fences, streams and a range of other linear features. Recorded linear features have a minimum length of 20 m and may include gaps of up to 20 m. All linear features are recorded unless they form part of a curtilage or they are within the woodland canopy. Woody linear features, including hedges, remnant hedges and lines of trees were classified using a key developed for CS in 2005–2006 (Maskell et al., 2008) following consultation with the Hedgerow Steering Group of the UK BAP. Precise definitions of features were recorded in the field handbooks (Barr, 1990, 1998; Maskell et al., 2008). Linear features were not recorded in detail in 1978 but were recorded subsequently in 1984, 1990, 1998 and 2007.

3.3 Point features

Point features are individual landscape elements that occupy less than an area of 20 km × 20 m. Point features may be trees or groups of trees, ponds and other freshwater features, physiological features such as cliffs, buildings and other structures with various use codes (for example, “residential” or “agricultural”). As with lines and areas, points are drawn and recorded on base maps using standard codes (Maskell et al., 2008).

4 Data collected

4.1 Area data

A summary of the categories of area data collected is given in Table 1. The areas of polygons allocated to broad habitats are available for each survey year. Additional information was collected in 1984, 1990, 1998 and 2007. This additional information includes the broad land use category of each polygon, a list of key species in the polygon and cover,

Table 1. Data collected within mapped polygons.

Attribute	Description	1978	1984	1990	1998	2007
Broad/priority habitat area	BAP broad habitat	×	×	×	×	×
Theme	Broad land use category, e.g. "agricultural crops"	(×)	×	×	×	×
Primary attribute	Feature name, e.g. "potatoes"	(×)	×	×	×	×
Species	Species where relevant	(×)	×	×	×	×
Species cover	Cover of above species across polygon		×	×	×	×
Primary qualifier	Additional information pertaining to primary attribute		×	×	×	×
Structure use	Use, where theme is "structures"		×	×	×	×
Physiography cover	Cover, used where theme is "inland physiography"		×	×	×	×
Road verge A	Width of verge A where theme is "transport"		×	×	×	×
Road verge B	Width of verge B where theme is "transport"		×	×	×	×
Modal DBH	Modal diameter at breast height (DBH), where theme is "forestry"		×	×	×	×
Mosaic percent area	If broad habitat is classed as mosaic, % cover of each primary attribute				×	×

diameter at breast height (DBH) of trees where the polygon is forestry, and the width of verges where the element is a "transport" type (such as a road).

4.2 Linear feature data

A summary of the categories of linear feature data collected is given in Table 2. Descriptions of features and attributes are available for each survey year between 1984 and 2007, with additional detail being collected in 1998 and 2007 regarding the condition of hedgerows, such as widths, signs of management, and margins.

4.3 Point feature data

A summary of the categories of point feature data collected is given in Table 3. Data on the type of feature have been collected in each survey since 1984, including details on species, use (where appropriate) and DBH of trees. Additional information regarding veteran trees was recorded in 1998 and 2007, when up to 10 veteran trees were recorded per square, consisting of the first two veteran trees of each species encountered in the field.

5 Data quality

Each field survey was carried out by teams of experienced botanical surveyors, and was preceded by an intensive training course, ensuring high standards and consistency of methodology, identification, effort and recording across CS according to criteria laid out in the field handbooks (Barr, 1984, 1990, 1998; Bunce, 1978; Maskell et al., 2008). During the surveys, survey teams were initially supervised and later monitored by experienced project staff.

Data were recorded on waterproof paper sheets in 1978, 1984, 1990 and 1998 and were subsequently digitised from the field sheets, following defined procedures. The digitised data have always been stored in secure, regularly backed-up databases. The 1984, 1990 and 1998 data were digitised in the 1990s, the linework being stored in Esri's ArcINFO™ geographical information system (Esri, 2017) coverages with the attributes being stored in an Oracle (Oracle Corporation, 2017) database. Before the 2007 survey, a data migration process was undertaken to transform each survey's data set into matching schemas, incorporating the point and linework and attributes into a geodatabase stored in Oracle and accessed via ArcSDE (Esri, 2017). The habitat polygons from 1978 were not digitised until 2009 (Wood et al., 2012), and were thus not reported in the main report for CS 2007 (Carey et al., 2008).

The move to electronic capture methods using a specially designed software package ("CS Surveyor") in 2007 removed the need for post-survey digitising and therefore eliminated a potential source of error. Improvements to data quality resulted from the inclusion of mandatory data entry fields for each feature, prompts for expected data for each of the mapped feature types and the removal of issues of illegible records. The use of a digital system enabled surveyors to ensure that each of the mapped components had been visited and to record whether change had occurred against each entry. This requirement to record change was a compulsory element of the survey, enforced by the digital system when any changes were made to mapped habitats and features. Additionally, the data were transferred back to the office soon after completion, enabling prompt data checking. Surveyors and managers could communicate readily to discuss any issues arising.

Surveyors had the ability to improve the quality of data from a prior survey by "back-allocating" features if they thought they had been recorded incorrectly in the previous survey. For example, if they encountered a large oak tree that had not been previously recorded, they would know that such

Table 2. Data collected regarding linear features.

Attribute	Description	1984	1990	1998	2007
Length	Length of feature	×	×	×	×
Theme	Feature name, e.g. "bank", "inland water", "woody linear feature"	×	×	×	×
Primary attribute	Feature type, e.g. "stone bank", "canal"	×	×	×	×
Height	Height of feature, where appropriate	×	×	×	×
Base height	Basal height of feature (hedgerow)			×	×
Width	Width of feature (hedgerow)			×	×
Modal DBH	Diameter at breast height (DBH), where appropriate	×	×	×	×
Condition	Condition assessment (walls, fences)	×	×	×	×
Historic management	Evidence of historic management (hedgerow)			×	×
Evidence management	Evidence of recent management (none, newly planted, cutting e.g. flail or saw (< 3 years), laying or coppicing) (< 5 years), both of the preceding two		×	×	×
Staked trees	Staked individual trees within the feature (hedgerow)		×	×	×
Tree protectors	Tree protectors		×	×	×
Line of stumps	Whether feature is a line of stumps (hedgerow)		×	×	×
Vertical gappiness	% of breaks which extend from canopy to ground along hedgerow			×	×
Margin width left	Margin width on left side of feature			×	×
Margin width right	Margin width on right side of feature			×	×
Species composition	Mixed species, > 50 % hawthorn, > 50 % other (hedgerow)	×	×	×	×
Species	Tree/shrub species (hedgerow)	×	×	×	×
Proportion	Proportion of species in feature (hedgerow)	×	×	×	×

Table 3. Data collected regarding point features.

	Attribute	Description	1984	1990	1998	2007
	Theme	Broad land use category e.g. "forestry", "building"	×	×	×	×
	Primary attribute	Feature type, e.g. "individual tree"	×	×	×	×
	Species	Species, where relevant	×	×	×	×
	Proportion	Proportion of species in feature	×	×	×	×
	Use	Use where appropriate, e.g. "agricultural"	×	×	×	×
	Buffer	Buffer zone present			×	×
	Modal DBH	Modal diameter at breast height (DBH)	×	×	×	×
Veteran trees	Tree dead	Dead tree			×	×
	Missing limbs	Missing branches			×	×
	Dead wood	Dead wood attached to trunk			×	×
	Dead missing bark	Dead, loose missing bark			×	×
	Lightning strikes	Evidence of lightning strikes			×	×
	Hollow trunk	Hollow trunk or major rot sites			×	×
	Veteran tree type	Standard, pollard or layered			×	×
	Epiphyte cover	Epiphytes: rare, present, abundant?			×	×
	Ivy cover	Ivy cover: 30 or > 30 %			×	×
	Canopy live	% of canopy live			×	×

a feature could not have grown in the few intervening years between surveys, and hence could confidently record that the feature must have been in existence in the previous survey, but must have missed being recorded. In this case, the information for the feature (point, line or area) would be back-allocated in the database for the previous survey.

Quality assurance (QA) exercises were undertaken during the 1990, 1998 and 2007 surveys, which involved a second team of surveyors (QA assessors) repeating the survey for all or part of a square. The 1990 QA report (Prosser and Wallace, 1992) cited an 89 % agreement between the field surveyors and QA assessors recording of primary land cover codes, and an 80 % agreement for primary boundary codes. The 1998 report (Prosser and Wallace, 1999) gave an 88 % agreement between the field surveyors and QA assessors recording of primary land cover codes, and an 85 % agreement for primary boundary codes. In 2007 (Norton et al., 2008), an assessment was made of point features, which had an 89 % agreement, and linear features, which had a 99 % match and a 73 % agreement at the polygon level. This last figure appears lower than the figures in the 1990 and 1998 assessments, but was assessed slightly differently, being at the broad/priority habitat level, rather than the primary code level. This introduced minor discrepancies (particularly between the choice of broad or priority habitat) which could largely be rectified with post-processing of the data.

A limitation of the datasets is that the exact site locations are held confidentially to protect landowner privacy (most sites are privately owned and surveys are only undertaken with prior permission), and also to secure the long-term nature of the project. As a consequence of this, the raw habitat data are not available as spatial datasets but rather as flat files, which may be analysed spatially at the level of the 45 environmental land classes, in conjunction with the ITE Land Classification dataset (Bunce et al., 2007). Regional estimates below the level of these land classes are not statistically robust due to sampling limitations, as described by Bunce et al. (1996b). National estimates are available in spatial formats, as detailed in Sect. 10. Within these national estimate datasets, statistical upper and lower limits are provided for each feature in question.

6 Methodological development

The success of the sampling methodology overall has been discussed in Wood and Bunce (2016). The method of habitat mapping is deemed to be highly successful for collecting the necessary data, and it is currently the only method that provides such detailed information at a national level, with the additional benefit of being able to assess change reliably. Where sufficient resource is available, changes to the methodology would not be recommended. However, although CS is a sample survey, field data collection is still a relatively expensive method of gathering information and various other options for capturing the same information have been proposed, particularly using remotely sensed methods. An assessment of using aerial photography to map habitats (in particular, condition) has been made in Wood et al. (2015). Broad habitats can be generally mapped from aerial photography, with some habitats, such as broadleaved woodland, being more successfully mapped from the air than others, such as fen, marsh and swamp. However, many elements, especially structural and species attributes, cannot be mapped successfully from the air. No detailed measurements or condition assessments were possible for any landscape feature from the aerial photos alone. Virtually no species were identified for most of the feature types (although infrared photography has now been used to improve habitat identification and composition; Ståhl et al., 2011). Several broad habitats rely on a thorough knowledge of the plant species occurring there, before a correct identification can be made. This is particularly important in differentiating between certain habitats, such as types of grassland and for the identification of priority habitats in particular (for example Purple Moor Grass Rush Pasture). In a survey such as CS, from which estimates for the whole of Great Britain are produced (Carey et al., 2008), a significant national underestimate of many features would result from mapping undertaken from aerial photos, and changes would be difficult, if not impossible, to assess. Similarly, although imagery available from drones is becoming more widespread, the detail remains below the level gained from field survey and whilst it could potentially be used to increase the speed and accuracy of mapping habitats and land cover extents, this would be at a cost.

These issues are also to be taken into account when assessing the use of satellite derived data, such as the series of land cover maps of Great Britain (Morton et al., 2011). Whilst these products give an excellent coverage of broad habitat extents for the whole of the United Kingdom, the ecological detail outlined above is not included, neither are details regarding point or line features (Fuller et al., 1994b). It is also not currently possible to estimate change from the land cover map series, as the earlier maps use different mapping classes to the later ones (classes that are not directly comparable), and in comparison with CS field data, they have an accuracy of approximately 62 % (in 2007) at the broad habitat level (Morton et al., 2011).

Rather than mapping the full extent of habitats and features within the 1 km squares, information based on a grid or dispersed points could potentially save time in the field (as, for example, in a 1990s survey of "key habitats"; Hornung et al., 1997; Barr et al., 2017). Whilst this would still provide the potential to produce national estimates for areas, much information regarding point and line features, and landscape structure and pattern would be lost. It is important that any new technologies or methodologies employed must be compatible with the existing databases.

Whilst certain concessions in recording could potentially be made in order to save resource, perhaps particularly for features/habitats with slow rates of change, the current methodology gives an optimal dataset for the full exploration of ecological issues particularly in relation to habitat change, some of which are outlined below.

7 Use of the data

The Countryside Survey provides a valuable resource, offering potential for a wide range of analyses at different temporal and spatial scales. A major benefit of the programme is the co-location of a wide range of recorded ecological variables (i.e. soil, vegetation, land cover and water). Monitoring of these variables is of key importance for identifying environmental change, evaluating policy responses and identifying drivers and processes of ecological change.

The results presented constitute the main findings from CS2007 that have, to date, appeared across a number of UK and country level reports for policy makers (countrysidesurvey.org.uk). CS, in common with comparable national surveys (e.g. Stahl et al., 2011), has been funded for both science and policy objectives (Norton et al., 2012b).

7.1 Stock and change: national estimates of broad habitat areas

The recording framework for broad habitats within CS makes it possible to provide national estimates for both the extent (in each survey year) and the change in extent (1990, 1998 and 2007) for broad and priority habitats, using the data from the Countryside Surveys. Estimates of change can be also made using the 1978 and 1984 data, but without the same level of confidence, due to the smaller sample size in those surveys. National estimates of the extent of 17 broad habitats and 12 priority habitats in 1998 and 2007 are presented in Carey et al. (2008). Priority habitats include upland mixed ash wood, wet woodland, upland oakwood, lowland mixed deciduous, upland birch woods, upland and lowland calcareous grassland, upland and lowland dwarf shrub heath, reedbed and purple moor grass rush pasture, as well as the linear feature priority habitat, hedges.

The condition of the vegetation surveyed in each broad habitat has been reported for the 1990, 1998 and 2007 Countryside Surveys. This is because the position of each vegetation plot is known (Wood et al., 2017) and so the species data recorded in each plot can be referenced to a specific broad habitat.

National estimates are based on calculations of the extents of each broad or priority habitat for each of the 45 land classes for England, Scotland and Wales individually, as well as for Great Britain. The procedure traditionally (up until 1998) used for calculating regional or national estimates was to produce means and standard errors for the quantity of interest for each Land Class and then to combine these to produce an estimated mean or total (with associated standard error) for the specified region as described by Haines-Young et al. (2003). The method of combination differed depending on whether a total or mean figure is required, but in both cases it involved weighting the individual land class estimates by values proportional to the area of land within the Land Class. Testing for significance requires more information about the distribution of an estimate than just its standard error. Prior to 1998, significance was assessed by assuming normality of estimates. In 1998, because of concerns about the validity of this assumption, largely because of the skewness of some of the features being estimated, standard errors and confidence intervals for square level data were estimated using the bootstrap method (Efron and Tibshirani, 1993). Essentially bootstrapping involves treating sample data as a population from which to resample. Each resample produces a separate estimate of some quantity of interest, for example stock or change. A large number of resamples (typically 1000 or 10 000) then gives an approximation to the distribution of the required estimate, from which any statistic can be extracted. The main advantage of this method of estimation for CS is that it allows for non-normality in the data, without requiring details of the actual distribution. As such it provides more accurate measurements of significance.

Prior to 2007, comparison between years was difficult because of the gradual increase in the number of sample squares in each of the years. In CS1998, the change between 1990 and 1998 was calculated using only data from those squares that were surveyed in both 1990 and 1998. Change was only calculated between squares that had been surveyed in both of the years in question, leading to minor discrepancies between the difference between the stock estimates, and the change estimates reported. In order to address this issue of incompatibility and to make better use of all the data collected in the survey, a new analytical procedure, the "consistent model", was developed for CS2007 which uses all available information from the time series (Scott, 2008). National estimates of broad habitats for each survey year for Great Britain are presented in Table 4.

It is important to note that the estimates for 1978 are not directly comparable to those for later surveys published in the Countryside Survey report for 2007 (Carey et al., 2008). This is primarily for two reasons. Firstly, due to the limited sample size of 256 1 km survey squares, estimates have been calculated using the 1990 ITE Land Classification (with 32 classes) rather than the revised 2007 Land Classification (with 45 classes) (see Sect. 2), as there are statistically not enough sample survey squares per class with 45 classes. (For work concentrating specifically on change, it is possible to calculate national estimates for all years using the 1990 Land Classification, but for GB as a whole only.) Secondly, due to the way broad habitats have been allocated retrospectively, habitats may not necessarily equate directly to the later datasets. The national estimates are publicly available, in addition to the raw data (Barr et al., 2014a, b, e, f, i, l, m; 2015a, b; Brown et al., 2014a, d; Bunce et al., 2012a, b).

7.2 Stock and change: national estimates of linear and point features

As with areas, the methods of recording linear and point features have been refined over time, but where there has been consistency of recording over time, the length of linear landscape features and the numbers of point features including trees and ponds (and changes in those lengths and numbers) can be estimated. Assessments of the condition of linear features are confined largely to more recent Countryside Surveys, in particular 1998 and 2007.

Linear features in the countryside are often complex and made up of different components; for example, a single field

Table 4. Estimated area (thousands of hectares) and percentage of land area of broad habitats in Great Britain from 1978 to 2007. Note that because of changes in definitions that have been applied retrospectively, the estimates from 1990 and more especially 1984 and 1978 are not in all cases directly comparable with later surveys. (Please note that not all the totals are equal to the sum of the column due to unavailable data.)

	Great Britain									
	1978		1984		1990		1998		2007	
Broad habitat	1000s ha	% area of GB	1000s ha	% area of GB	1000s ha	% area of GB	1000s ha	% area of GB	1000s ha	% area of GB
Broadleaved mixed and yew woodland	995	4.3	1317	5.6	1343	5.8	1328	5.7	1406	6.0
Coniferous woodland	1413	6.1	1243	5.3	1239	5.3	1386	5.9	1319	5.7
Boundary and linear features	364	1.6	491	2.1	581	2.5	511	2.2	496	2.1
Arable and horticulture	5105	21.9	5283	22.7	5025	21.6	5067	21.7	4608	19.8
Improved grassland	5188	22.3	5903	25.3	4619	19.8	4251	18.2	4494	19.3
Neutral grassland	1442	6.2	467	2.0	1669	7.2	2007	8.6	2176	9.3
Calcareous grassland	53	0.2	75	0.3	78	0.3	61	0.3	57	0.2
Acid grassland	1786	7.7	1476	6.3	1821	7.8	1502	6.4	1589	6.8
Bracken	258	1.1	439	1.9	272	1.2	315	1.3	260	1.1
Dwarf shrub heath	1677	7.2	1388	6.0	1436	6.2	1299	5.6	1343	5.8
Fen, marsh, swamp	231	1	428	1.8	427	1.8	425	1.8	392	1.7
Bog	2004	8.6	2303	9.9	2050	8.8	2222	9.5	2232	9.6
Standing open waters and canals	360	1.5	284	1.2	200	0.9	196	0.8	204	0.9
Rivers and streams	75	0.3	70	0.3	70	0.3	65	0.3	58	0.2
Montane	NA	NA	41	0.2	NA	NA	41	0.2	42	0.2
Inland rock	190	0.8	38	0.2	76	0.3	111	0.5	101	0.4
Built-up areas and gardens	1441	6.2	1268	5.4	1266	5.4	1279	5.5	1323	5.7
Other land	249	1.1	NA	NA	659	2.8	762	3.3	731	3.1
Unsurveyed urban land	482	2.1	NA	NA	482	2.1	482	2.1	482	2.1
Total area	23 313		23 313		23 313		23 313		23 313	

NA = not available

Table 5. The length (thousands of kilometres) and change in length of boundary and linear features in Great Britain, from 1984 to 2007. Arrows denote significant change ($p < 0.05$) in the direction shown.

					Direction of significant changes		
	1984	1990	1998	2007	84–90	90–98	98–07
Hedges	624	506	508	477	↓		↓
Line of trees/shrubs/relict hedge	58	71	109	114		↑	↑
Line of trees/shrubs/relict hedge/fence	32	59	99	114	↑	↑	↑
Walls	198	173	176	174	↓		↓
Bank/grass strip	56	57	62	64			
Fence	571	644	653	664	↑		↑

boundary may contain a fence, a hedge and a bank. To simplify reporting of these features, a hierarchy of feature types was used to define any compound linear feature (Carey et al., 2008) with ecologically important features, hedges and lines of trees at the top of the hierarchy. National estimates for linear feature types (in thousands of kilometres) were achieved by calculating a mean length for each feature type for the sample squares within a land class and then multiplying this figure by the number of 1 km squares in the land class in a similar method to that described in Sect. 7.1 for areas (Scott et al., 2008). This calculation gives an estimate of the total length in the land class and subsequently, by summation, of all land classes (Table 5). National estimates of ponds and hedgerow tree numbers can be derived in the same way. The national estimates for linear features are publicly available with associated error terms, in addition to the raw data (Barr et al., 2014d, c, h, g, k, j; Brown et al., 2014b, c).

Table 6. Table of DOIs for landscape element data.

	Landscape area data	Landscape linear feature data	Landscape point feature data
1978	https://doi.org/10.5285/86c017ba-dc62-46f0-ad13-c862bf31740e (Bunce et al., 2016)		
1984	https://doi.org/10.5285/b656bb43-448d-4b2c-aade-7993aa243ea3 (Barr et al., 2016g)	https://doi.org/10.5285/a3f5665c-94b2-4c46-909e-a98be97857e5 (Barr et al., 2016f)	https://doi.org/10.5285/124b872e-036e-4dd3-8316-476b5f42c16e (Barr et al., 2016e)
1990	https://doi.org/10.5285/94f664e5-10f2-4655-bfe6-44d745f5dca7 (Barr et al., 2016h)	https://doi.org/10.5285/311daad4-bc8c-485a-bc8a-e0d054889219 (Barr et al., 2016i)	https://doi.org/10.5285/1481bc63-80d7-4d18-bcba-8804aa0a9e1b (Barr et al., 2016a)
1998	https://doi.org/10.5285/1e050028-5c55-42f4-a0ea-c895d827b824 (Barr et al., 2016b)	https://doi.org/10.5285/8aaf6f8c-c245-46bb-8a2a-f0db012b2643 (Barr et al., 2016d)	https://doi.org/10.5285/ed10944f-40c8-4913-b3f5-13c8e844e153 (Barr et al., 2016c)
2007	https://doi.org/10.5285/bf189c57-61eb-4339-a7b3-d2e81fdde28d (Brown et al., 2016c)	https://doi.org/10.5285/e1d31245-4c0a-4dee-b36c-b23f1a697f88 (Brown et al., 2016b)	https://doi.org/10.5285/55dc5fd7-d3f7-4440-b8a7-7187f8b0550b (Brown et al., 2016a)

8 Wider uses of data to date

The potential uses of these unique data sets are wide-ranging, and can be broadly divided into two groups: investigations of ecological drivers and process, and provision of evidence to policy makers. CS data give a national overview of changes in habitats (Haines-Young et al., 2003; Firbank et al., 2003; Norton et al., 2012b; Howard et al., 2003). During the production of the National Ecosystem Assessment (NEA) in Britain, CS data (with the inclusion of the Northern Ireland survey) made a vital contribution to our understanding of ecosystems across the UK (NEA UK, 2011). The CS datasets have increasingly been used in the area of ecosystem services and natural capital, for assessing the scale of the benefits that ecosystems provide (Norton et al., 2012a) and investigating the distribution and interdependencies of specific environmental variables (Henrys et al., 2015).

The datasets can also be used to identify and quantify the extent of a particular species. For example, when the *Chalara fraxinea* ash dieback disease came to prominence in the news in 2012, CS data were used to produce a national picture of ash trees, supplementing information from the Forestry Commission regarding ash in larger woodlands (Forestry Commission, 2012). Estimates of ash as hedgerow trees (Maskell et al., 2013a), within areas less than half a hectare (Maskell et al., 2013b) and individual trees (Maskell et al., 2013c), were drawn from CS data. Data have also been used to assess relationships between wider species richness (birds and plants) and habitat and landscape feature presence and extent (Rhodes et al., 2015; Smart et al., 2010).

Drivers of environmental change may be investigated, for example the effects of agricultural intensification (Petit et al., 2004a) and farming practices (Potter and Lobley, 1996) on habitat quality and extent.

CS data have contributed to the area of woodland research, examining the effects of landscape structure on specific species (Petit et al., 2004a, b; Kimberley et al., 2016). The loss of hedgerows has been a key concern since the end of the Second World War, and CS data have proved useful in determining the extent and nature of changes since 1984 (Barr et al., 1991; Barr and Gillespie, 2000; Petit et al., 2003; Norton et al., 2012b) and applying these to policy changes (Barr and Parr, 1994). CS data have contributed to determining policy, for example the Hedgerow Regulations (The Hedgerows Regulations, 1997). CS linear data have been incorporated into other data products, for example Scholefield et al. (2016a, b).

9 Data availability

The datasets have been assigned digital object identifiers as in Table 6.

The most recent (2007) Land Classification is available as Bunce et al. (2007). National estimate datasets are also available as both non-spatial flat files (Barr et al., 2014a, d, e, g, i, k, l; Brown et al., 2014b, d; Bunce et al., 2012a) and spatial national datasets (Bunce et al., 2012b; Brown et al., 2014c, a; Barr et al., 2014b, c, f, h, j, m, 2015a, b).

The datasets are available from the CEH Environmental Information Data Centre Catalogue (https://catalogue.ceh.ac.uk, last access: 10 April 2018). Datasets are provided under the terms of the Open Government Licence (http://eidchub.ceh.ac.uk/administration-folder/tools/ceh-standard-licence-texts/ceh-open-government-licence/plain, last access:

10 April 2018, http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/, last access: 10 April 2018). The metadata are stored in the ISO 19115 (2003) schema (International Organization for Standardization, 2015) in the UK Gemini 2.1 profile (UK GEMINI, 2015). Users of the datasets will find the following documents useful (supplied as supporting documentation with the datasets): Bunce (1978), Maskell et al. (2008), Barr and Wood (2011), and Barr (1984, 1990, 1998).

10 Conclusions

The ecological landscape element data recorded during the Countryside Survey of Great Britain are an invaluable national resource, which, over the years, has proved useful to a range of users, including the scientific community and national policy makers. The data are collected in a statistically robust and quality-controlled manner, follow standard, repeatable methods and cover wide temporal and spatial scales. The intention is that a repeat survey will be undertaken in the near future (and a sub-sample of plots have already been surveyed in the summer of 2016, mainly in Wales, largely as part of the Glastir Monitoring and Evaluation Programme; Emmett and GMEP team, 2014). As a decade has now passed since the most recent full survey, an addition to this long-term national resource is becoming increasingly timely, particularly in these current times of political, socio-economic and climatic change. The latest news regarding further surveys can be found on the website for the programme, www.countrysidesurvey.org.uk (last access: 10 April 2018).

Author contributions. CMW prepared the manuscript with significant contributions from all co-authors, and is the current database manager for the Land Use Research Group at CEH Lancaster. RGHB designed the sampling framework and survey strategy in 1978. RGHB, SMS, LCM, LRN and DCH have all been part of the Countryside Survey scientific co-ordination team for at least one survey, with WAS and PAH contributing statistical support, and SMW, MJB, RJS, RS and JWW providing technical and data management support.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank all the landowners who kindly gave permission to survey their holdings in the survey sample squares in 1978, 1984, 1990, 1998 and 2007. Without their cooperation and assistance the Countryside Survey would not exist. We also acknowledge and thank all the field surveyors involved in each field campaign.

The most recent Countryside Survey was funded largely by the Natural Environment Research Council, the Department for Environment Food and Rural Affairs, and the Scottish and Welsh governments.

We would like to thank an anonymous reviewer and Geoffrey Griffiths for valuable comments which improved the manuscript.

Edited by: David Carlson
Reviewed by: Geoffrey Griffiths and one anonymous referee

References

Bakkestuen, V., Erikstad, L., and Halvorsen, R.: Step-less models for regional environmental variation in Norway, J. Biogeogr., 35, 1906–1922, 2008.

Barr, C.: Countryside Survey 1984 field handbook, Institute of Terrestrial Ecology, Grange-over-Sands, 1984.

Barr, C.: Countryside Survey 1990 field handbook, Institute of Terrestrial Ecology, Grange-over-Sands, 1990.

Barr, C. and Gillespie, M.: Estimating hedgerow length and pattern characteristics in Great Britain using Countryside Survey data, J. Environ. Manage., 60, 23–32, https://doi.org/10.1006/jema.2000.0359, 2000.

Barr, C., Howard, D., Bunce, B., Gillespie, M., and Hallam, C.: Changes in hedgerows in Britain between 1984 and 1990. Contract report to the Department Of The Environment, Institute of Terrestrial Ecology, Grange-over-Sands, 1991.

Barr, C. J.: Countryside Survey 2000 Field Handbook, Institute of Terrestrial Ecology, Grange-over-Sands, 1998.

Barr, C. J. and Parr, T. W.: Hedgerows: Linking Ecological Research and Countryside Policy, in: Hedgerow Management and Nature Conservation, edited by: Watt, T. A. and Buckley, G. P., Wye College Press, University of London, 1994.

Barr, C. J. and Wood, C. M.: The Sampling Strategy for Countryside Survey (up to 2007). Revised and Updated from: "The Sampling Strategy for Countryside Survey", C. J. Barr, September 1998. DETR CONTRACT No. CR0212, Lancaster, 2011.

Barr, C. J., Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Clarke, R. T., Crowe, A., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Countryside Survey estimates of Broad Habitat area change in Great Britain between 1990, 1998 and 2007, NERC Environmental Information Data Centre, Lancaster, UK, https://doi.org/10.5285/7e2981e7-bd4c-4992-b7b0-1b1253bfd20d, 2014a.

Barr, C. J., Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Clarke, R. T., Crowe, A., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey mapped estimates of Broad Habitat area change in Great Britain between 1990 and 2007, NERC Environmental Information Data Centre, https://doi.org/10.5285/4af5abe4-158a-4736-b318-ec660e09e45a, 2014b.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 1990 mapped estimates of linear feature lengths in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/ce514508-bd25-403e-bb35-42001e3c2b25, 2014c.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 1990 estimates of linear feature lengths in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/ebbd1ba0-4497-4c80-891a-34ef3b8989cd, 2014d.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 1990 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/32dda22c-24cc-4fa8-adcb-5481b1b77ef1, 2014e.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 1990 mapped estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/53ef00f4-e0c5-4095-850e-d4c474bc4ffd, 2014f.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Countryside Survey 1998 estimates of linear feature lengths in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/59422890-c8d3-49b8-b17e-4f5062475140, 2014g.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Countryside Survey 1998 mapped estimates of linear feature lengths in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/6bc0ccb4-63c8-4121-bc4b-3c479877cff9, 2014h.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Countryside Survey 1998 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/acbb0f9c-f2a5-4cbd-97c3-84e786f2e479, 2014i.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., and Wright, S. M.: Countryside Survey 1984 mapped estimates of linear feature lengths in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/362bff26-d2ca-4700-b1f0-22072f4bc0cc, 2014j.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., and Wright, S. M.: Countryside Survey 1984 estimates of linear feature lengths in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/271d398c-29de-4ff9-ba68-1b8e0621c6ee, 2014k.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., and Wright, S. M.: Countryside Survey 1984 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/ad823866-cded-4778-9bd7-999c821b120e, 2014l.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., and Wright, S. M.: Countryside Survey 1984 mapped estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/5964c7a2-0537-40d7-9d1b-76c37c3aadf2, 2014m.

Barr, C. J., Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Countryside Survey mapped estimates of Broad Habitat area change in Great Britain between 1998 and 2007, NERC Environmental Information Data Centre, https://doi.org/10.5285/d83a0f9e-00c9-4d2d-9d0a-e92a16dcb334, 2015a.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Countryside Survey mapped estimates of Broad Habitat area change in Great Britain between 1990 and 1998, NERC Environmental Information Data Centre, https://doi.org/10.5285/2bfdede9-8008-4ba3-ac8e-af4e6ab9888b, 2015b.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Landscape point feature data 1990 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/1481bc63-80d7-4d18-bcba-8804aa0a9e1b, 2016a.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Landscape area data 1998 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/1e050028-5c55-42f4-a0ea-c895d827b824, 2016b.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Landscape point feature data 1998 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/ed10944f-40c8-4913-b3f5-13c8e844e153, 2016c.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Watkins, J. W., Wood, C. M., and Wright, S. M.: Landscape linear feature data 1998 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/8aaf6f8c-c245-46bb-8a2a-f0db012b2643, 2016d.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., and Wright, S. M.: Landscape point feature data 1984 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/124b872e-036e-4dd3-8316-476b5f42c16e, 2016e.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., and Wright, S. M.: Landscape linear feature data 1984 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/a3f5665c-94b2-4c46-909e-a98be97857e5, 2016f.

Barr, C. J., Bunce, R. G. H., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Whittaker, H. A., Wood, C. M., and Wright, S. M.: Landscape area data 1984 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/b656bb43-448d-4b2c-aade-7993aa243ea3, 2016g.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Landscape area data 1990 [Countryside Survey], https://doi.org/10.5285/94f664e5-10f2-4655-bfe6-44d745f5dca7, 2016h.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Landscape linear feature data 1990 [Countryside Survey], 2016i.

Barr, C. J., Bunce, R. G. H., Cummins, R. P., Hallam, C. J., Hornung, M., and Wood, C. M.: Habitat and vegetation data from an ecological survey of terrestrial key habitats in England, 1992–1993, NERC Environmental Information Data Centre, https://doi.org/10.5285/7aefe6aa-0760-4b6d-9473-fad8b960abd4, 2017.

Bartholomé, E. and Belward, A.: GLC2000: a new approach to global land cover mapping from Earth observation data, Int. J. Remote Sens., 26, 1959–1977, https://doi.org/10.1080/01431160412331291297, 2005.

Brandt, J., Bunce, R., Howard, D., and Petit, S.: General principles of monitoring land cover change based on two case studies in Britain and Denmark, Landscape Urban Plan., 62, 37–51, 2002.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 2007 mapped estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/bb785c19-3cf6-4483-9fe5-13829d489a83, 2014a.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 2007 estimates of linear feature lengths in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/e687330b-a0f7-45a1-b58c-398e67da3028, 2014b.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 2007 mapped estimates of linear feature lengths in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/fc65177d-b113-420e-a70b-05d3f42682d5, 2014c.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Countryside Survey 2007 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/f03cba75-8bca-4679-ae2a-77a9fcbd4df3, 2014d.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chamberlain, P. M., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Landscape point feature data 2007 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/55dc5fd7-d3f7-4440-b8a7-7187f8b0550b, 2016a.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Landscape linear feature data 2007 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/e1d31245-4c0a-4dee-b36c-b23f1a697f88, 2016b.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., and Wright, S. M.: Landscape area data 2007 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/bf189c57-61eb-4339-a7b3-d2e81fdde28d, 2016c.

Bunce, R. G. H.: UK Ecological Survey. Handbook of Field Methods, Institute of Terrestrial Ecology, Grange-over-Sands, 1978.

Bunce, R. G. H. and Smith, R. S.: An ecological survey of Cumbria, Structure Plan Working Paper, 4, Cumbria County Council and Lake District Special Planning Board, Kendal, 1978.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE Land Classification of Great Britain 1990, NERC Environmental Information Data Centre, https://doi.org/10.5285/ab320e08-faf5-48e1-9ec9-77a213d2907f, 1990.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE Merlewood Land Classification of Great Britain, J. Biogeogr., 23, 625–634, https://doi.org/10.1111/j.1365-2699.1996.tb00023.x, 1996a.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: Land classification for strategic ecological survey, J. Environ. Manage., 47, 37–60, https://doi.org/10.1006/jema.1996.0034, 1996b.

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., and Howard, D. C.: The ITE Land Classification: providing an environmental stratification of Great Britain, Environ. Monit. Assess., 39, 39–46, https://doi.org/10.1007/978-94-009-1653-1_5, 1996c.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D., and Scott, A.: ITE Land Classification of Great Britain 2007, NERC Environmental Information Data Centre, https://doi.org/10.5285/5f0605e4-aa2a-48ab-b47c-bf5510823e8f, 2007.

Bunce, R. G. H., Wood, C. M., Henrys, P. A., Smart, S. M., Howard, D. C., and Barr, C. J.: Countryside Survey 1978 estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/c8feabac-7c3f-4c35-91f7-296782b957d1, 2012a.

Bunce, R. G. H., Wood, C. M., Henrys, P. A., Smart, S. M., Howard, D. C., and Barr, C. J.: Countryside Survey 1978 mapped estimates of Broad Habitat areas in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/df180b2e-36ea-410e-8487-5a09942afa9e, 2012b.

Bunce, R. G. H., Bogers, M. M. B., Evans, D., and Jongman, R. H. G.: Field identification of habitats directive Annex I habitats as a major European biodiversity indicator, Ecological Indicators, 33, 105–110, https://doi.org/10.1016/j.ecolind.2012.10.004, 2013.

Bunce, R. G. H., Wood, C. M., Henrys, P. A., Smart, S. M., Howard, D. C., and Barr, C. J.: Landscape area data 1978 [Countryside Survey], NERC Environmental Information Data Centre, https://doi.org/10.5285/86c017ba-dc62-46f0-ad13-c862bf31740e, 2016.

Carey, P. D., Wallis, S., Chamberlain, P. M., Cooper, A., Emmett, B. A., Maskell, L. C., McCann, T., Murphy, J., Norton, L. R., Reynolds, B., Scott, W. A., Simpson, I. C., Smart, S. M., and Ullyett, J. M.: Countryside Survey: UK Results from 2007, NERC/Centre for Ecology & Hydrology, Lancaster, 2008.

Cole, B., King, S., Ogutu, B., Palmer, D., Smith, G., and Balzter, H.: Corine land cover 2012 for the UK, Jersey and Guernsey, NERC Environmental Information Data Centre, https://doi.org/10.5285/32533dd6-7c1b-43e1-b892-e80d61a5ea1d, 2015.

Cooper, A.: Land cover monitoring in Northern Ireland, Biodiversity: the Irish dimension, edited by: Rushton, B. S., Royal Irish Academy, Dublin, 122–131, 2000

Cooper, A., McCann, T., and Rogers, D.: Northern Ireland countryside survey 2007: Broad Habitat change 1998–2007, Northern Ireland Environment Agency, Coleraine, 2009.

Dunbar, M., Murphy, J., Clarke, R., Baker, R., Davies, C., and Scarlett, P.: Countryside Survey: headwater streams report from 2007, CS Technical Report No. 8/07, NERC Centre for Ecology and Hydrology, Wallingford, 2010.

Efron, B. and Tibshirani, R. J.: An introduction to the bootstrap, Chapman and Hall, London, 1993.

Elena-Rosselló, R.: Biogeoclimatic classification of the Spanish regions of Iberian Peninsula and Balearic islands, Ministry of Agriculture, Fisheries and Food, Madrid, 2007.

Emmett, B. E. and GMEP team: Glastir Monitoring & Evaluation Programme, First Year Annual Report to Welsh Government (Contract reference: C147/2010/11), NERC/Centre for Ecology & Hydrology, Bangor, 442 pp., 2014.

Emmett, B. A., Reynolds, B., Chamberlain, P. M., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A. J., Poskitt, J., Potter, E., Robinson, D. A., Scott, A., Wood, C., and Woods, C.: Countryside Survey: Soils Report from 2007, Centre for Ecology & Hydrology, Wallingford, UK, 2010.

Esri: http://www.esri.com/, last access: 23 August 2017.

Eva, H. D., Belward, A. S., De Miranda, E. E., Di Bella, C. M., Gond, V., Huber, O., Jones, S., Sgrenzaroli, M., and Fritz, S.: A land cover map of South America, Glob. Change Biol., 10, 731–744, https://doi.org/10.1111/j.1529-8817.2003.00774.x, 2004.

Firbank, L. G., Barr, C. J., Bunce, R. G. H., Furse, M. T., Haines-Young, R. H., Hornung, M., Howard, D. C., Sheail, J., Sier, A. R. J., and Smart, S. M.: Assessing stock and change in land cover and biodiversity in GB: an introduction to the Countryside Survey 2000, J. Environ. Manage., 67, 207–218, https://doi.org/10.1016/S0301-4797(02)00174-3, 2003.

Forestry Commission: NFI preliminary estimates of quantities of broadleaved species in British woodlands, with special focus on ash, Forestry Commission, available at: https://www.forestry.gov.uk/pdf/NFI_Prelim_BL_Ash_Estimates.pdf/$FILE/NFI_Prelim_BL_Ash_Estimates.pdf (last access: 10 April 2018), Edinburgh, 2012.

Fuller, R. M., Groom, G. B., and Jones, A. R.: The land cover map of Great Britain: an automated classification of Landsat Thematic Mapper data, Photogramm. Eng. Rem. S., 60, 553–562, 1994a.

Fuller, R. M., Sheail, J., and Barr, C. J.: The Land of Britain, 1930–1990: a comparative study of Field Mapping and Remote Sensing Techniques, Geogr. J., 160, 173–184, https://doi.org/10.2307/3060075, 1994b.

Fuller, R. M., Smith, G. M., Sanderson, J. M., Hill, R. A., Thomson, A. G., Cox, R., Brown, N. J., Clarke, R. T., Rothery, P., and Gerard, F. F.: Land Cover Map 2000 – Module 7 final report, NERC Centre for Ecology & Hydrology, Huntingdon, 2001.

Haines-Young, R., Barr, C., Firbank, L., Furse, M., Howard, D., McGowan, G., Petit, S., Smart, S., and Watkins, J.: Changing landscapes, habitats and vegetation diversity across Great Britain, J. Environ. Manage., 67, 267–281, https://doi.org/10.1016/S0301-4797(02)00179-2, 2003.

Henrys, P., Bee, E., Watkins, J., Smith, N., and Griffiths, R.: Mapping natural capital: optimising the use of national scale datasets, Ecography, 38, 632–638, https://doi.org/10.1111/ecog.00402, 2015.

Hornung, M. F., Barr, C. J., and Bunce, R. G. H.: Current status and prospects for key habitats in England. Part 6: Summary report, Institute of Terrestrial Ecology, Grange-over-Sands, 1997.

Howard, D., Watkins, J., Clarke, R., Barnett, C., and Stark, G.: Estimating the extent and change in Broad Habitats in Great Britain, J. Environ. Manage., 67, 219–227, https://doi.org/10.1016/S0301-4797(02)00175-5, 2003.

International Organization for Standardization: ISO 19115:2003: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020, last access: 12 October 2015.

Jackson, D.: Guidance on the Interpretation of the Biodiversity Broad Habitat Classification (Terrestrial and Freshwater Types): Definitions and the Relationship with Other Habitat Classifications (JNCC Report, No 307), JNCC, published online, available at: http://jncc.defra.gov.uk/page-2433 (last access: 10 April 2018), 73 pp., 2000.

Kimberley, A., Blackburn, G. A., Whyatt, J. D., and Smart, S. M.: How well is current plant trait composition predicted by modern and historical forest spatial configuration?, Ecography, 39, 67–76, https://doi.org/10.1111/ecog.01607, 2016.

Maddock, A.: UK Biodiversity Action Plan; Priority Habitat Descriptions, published online, available at: http://jncc.defra.gov.uk/PDF/UKBAP_PriorityHabitatDesc-Rev2011.pdf (last access: 10 April 2018), 2008.

Maskell, L. C., Norton, L. R., Smart, S. M., Carey, P. D., Murphy, J., Chamberlain, P. M., Wood, C. M., Bunce, R. G. H., and Barr, C. J.: Countryside Survey. Field Mapping Handbook CS Technical Report No. 1/07, Centre for Ecology & Hydrology, Lancaster, 2008.

Maskell, L. C., Henrys, P. A., Norton, L. R., Smart, S. M., and Wood, C. M.: Distribution of Ash trees (Fraxinus excelsior) in Countryside Survey data, NERC Centre for Ecology & Hydrology, Lancaster, 2013a.

Maskell, L. C., Henrys, P. A., Norton, L. R., Smart, S. M., and Wood, C. M.: Distribution of ash trees within areas less than half a hectare in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/7f7c99c7-5457-444d-978a-fe7b01a85fd0, 2013b.

Maskell, L. C., Henrys, P. A., Norton, L. R., Smart, S. M., and Wood, C. M.: Distribution of individual ash trees in Great Britain, NERC Environmental Information Data Centre, https://doi.org/10.5285/0c3567a8-3700-4d52-a21f-de1bd709141a, 2013c.

Mayaux, P., Bartholomé, E., Fritz, S., and Belward, A.: A new land-cover map of Africa for the year 2000, J. Biogeogr., 31, 861–877, 2004.

Metzger, M. J., Brus, D. J., Bunce, R. G. H., Carey, P. D., Gonçalves, J., Honrado, J. P., Jongman, R. H. G., Trabucco, A., and Zomer, R.: Environmental stratifications as the basis for national, European and global ecological monitoring, Ecol. Indic., 33, 26–35, https://doi.org/10.1016/j.ecolind.2012.11.009, 2013.

Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R., and Simpson, I.: Final Report for LCM2007 – the new UK land cover map, CS Technical Report No. 9/07, NERC/Centre for Ecology & Hydrology, Wallingford, 2011.

NEA UK: The UK national ecosystem assessment, UNEP-WCMC, Cambridge, 2011.

Norton, L. R., Scholefield, P. S., Maskell, L. C., Smart, S. M., Murphy, J., Williams, P., Scarlett, P., and Clarke, R.: Quality Assurance Report: MAPPING QUALITY ASSURANCE EXERCISE Centre for Ecology and Hydrology, Lancaster, 2008.

Norton, L., Inwood, H., Crowe, A., and Baker, A.: Trialling a method to quantify the "cultural services" of the English landscape using Countryside Survey data, Land Use Policy, 29, 449–455, https://doi.org/10.1016/j.landusepol.2011.09.002, 2012a.

Norton, L. R., Maskell, L. C., Smart, S. M., Dunbar, M. J., Emmett, B. A., Carey, P. D., Williams, P., Crowe, A., Chandler, K., Scott, W. A., and Wood, C. M.: Measuring stock and change in the GB countryside for policy: key findings and developments from the Countryside Survey 2007 field survey, J. Environ. Manage., 113, 117–127, https://doi.org/10.1016/j.jenvman.2012.07.030, 2012b.

Oracle: https://www.oracle.com, last access: 23 August 2017.

Petit, S., Stuart, R., Gillespie, M., and Barr, C.: Field boundaries in Great Britain: stock and change between 1984, 1990 and 1998, J. Environ. Manage., 67, 229–238, https://doi.org/10.1016/S0301-4797(02)00176-7, 2003.

Petit, S., Griffiths, L., Smart, S. S., Smith, G. M., Stuart, R. C., and Wright, S. M.: Effects of area and isolation of woodland patches on herbaceous plant species richness across Great Britain, Landsc. Ecol., 19, 463–471, https://doi.org/10.1023/B:LAND.0000036141.30359.53, 2004a.

Petit, S., Howard, D. C., and Stuart, R. C.: A national perspective on recent changes in the spatial characteristics of woodland in the British landscape, Landscape Urban Plan., 69, 127–135, https://doi.org/10.1016/j.landurbplan.2003.10.031, 2004b.

Potter, C. and Lobley, M.: The farm family life cycle, succession paths and environmental change in Britain's countryside, J. Agr. Econ., 47, 172–190, https://doi.org/10.1111/j.1477-9552.1996.tb00683.x, 1996.

Prosser, M. V. and Wallace, H. L.: Countryside Survey 1990 Quality Assurance Exercise, Report to Institute of Terrestrial Ecology, Merlewood Research Station, Canon Pyon, Herefordshire, 1992.

Prosser, M. V. and Wallace, H. L.: Countryside Survey 2000 Quality Assurance Exercise, Report to Institute of Terrestrial Ecology, Merlewood Research Station, Second draft, Canon Pyon, Herefordshire, 1999.

Rhodes, C. J., Henrys, P., Siriwardena, G. M., Whittingham, M. J., and Norton, L. R.: The relative value of field survey and remote sensing for biodiversity assessment, Methods Ecol. Evol., 6, 772–781, https://doi.org/10.1111/2041-210X.12385, 2015.

Romão, C.: Interpretation manual of European Union habitats, version EUR 28, available at: http://ec.europa.eu/environment/nature/legislation/habitatsdirective/docs/Int_Manual_EU28.pdf (last access: 10 April 2018), 2013.

Scholefield, P., Morton, D., Rowland, C., Henrys, P., Howard, D., and Norton, L.: A model of the extent and distribution of woody linear features in rural Great Britain, Ecol. Evol., 6, 8893–8902, https://doi.org/10.1002/ece3.2607, 2016a.

Scholefield, P. A., Morton, R. D., Rowland, C. S., Henrys, P. A., Howard, D. C., and Norton, L. R.: Woody linear features framework, Great Britain v.1.0, NERC Environmental Information Data Centre, https://doi.org/10.5285/d7da6cb9-104b-4dbc-b709-c1f7ba94fb16, 2016b.

Scott, W. A.: Countryside Survey. Statistical Report, Countryside Survey Technical Report No. 4/07, NERC Centre for Ecology and Hydrology, Lancaster, 2008.

Sheail, J. and Bunce, R. G. H.: The development and scientific principles of an environmental classification for strategic ecological survey in the United Kingdom, Environ. Conserv., 30, 147–159, https://doi.org/10.1017/S0376892903000134, 2003.

Smart, S., Maskell, L., Dunbar, M., Emmett, B., Marks, S., Norton, L., Rose, P., and Simpson, I.: An Integrated Assessment of Countryside Survey data to investigate Ecosystem Services in Great Britain. Countryside Survey Technical Report No. 10/07, NERC Centre for Ecology and Hydrology, Lancaster, 2010.

Ståhl, G., Allard, A., Esseen, P.-A., Glimskär, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Torell, Å. G., and Högström, M.: National Inventory of Landscapes in Sweden (NILS) – scope, design, and experiences from establishing a multiscale biodiversity monitoring system, Environ. Monit. Assess., 173, 579–595, https://doi.org/10.1007/s10661-010-1406-7, 2011.

The Hedgerows Regulations: The Hedgerows Regulations 1997, available at: http://www.legislation.gov.uk/uksi/1997/1160/contents/made (last access 10 April 2018), 1997.

UK GEMINI: UK GEMINI, http://www.agi.org.uk/join-us/agi-groups/standards-committee/uk-gemini, last access: 12 May 2015.

Villoslada, M., Bunce, R. G. H., Sepp, K., Jongman, R. H. G., Metzger, M. J., Kull, T., Raet, J., Kuusemets, V., Kull, A., and Leito, A.: A framework for habitat monitoring and climate change modelling: construction and validation of the Environmental Stratification of Estonia, Reg. Environ. Change, 17, 1–15, https://doi.org/10.1007/s10113-016-1002-7, 2016.

Williams, P., Biggs, J., Crowe, A., Murphy, J., Nicolet, P., Weatherby, A., and Dunbar, M.: Countryside survey: ponds report from 2007. CS Technical Report No. 7/07. Pond Conservation and NERC/Centre for Ecology & Hydrology, Wallingford, 2010.

Wood, C. M. and Bunce, R. G. H.: Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring, Earth Syst. Sci. Data, 8, 89–103, https://doi.org/10.5194/essd-8-89-2016, 2016.

Wood, C. M., Howard, D. C., Henrys, P. A., Bunce, R. G. H., and Smart, S. M.: Countryside Survey: measuring habitat change over 30 years: 1978 data rescue – final report, NERC Centre for Ecology & Hydrology, Lancaster, 2012.

Wood, C. M., Norton, L. R., and Rowland, C. S.: What are the costs and benefits of using aerial photography to survey habitats in 1 km squares?, NERC Centre for Ecology & Hydrology, Lancaster, UK, 2015.

Wood, C. M., Smart, S. M., Bunce, R. G. H., Norton, L. R., Maskell, L. C., Howard, D. C., Scott, W. A., and Henrys, P. A.: Long-term vegetation monitoring in Great Britain – the Countryside Survey 1978–2007 and beyond, Earth Syst. Sci. Data, 9, 445–459, https://doi.org/10.5194/essd-9-445-2017, 2017.

Wyatt, B., Greatorex-Davies, J., Hill, M., Parr, T., Bunce, R., and Fuller, R.: Comparisons of land cover definitions, Department of Environment Countryside 1990 Series, Vol. 3, HMSO, London, 1994.

Earth Syst. Sci. Data, 9, 445–459, 2017
https://doi.org/10.5194/essd-9-445-2017

Open Access Earth System Science Data

Long-term vegetation monitoring in Great Britain – the Countryside Survey 1978–2007 and beyond

Claire M. Wood[1], Simon M. Smart[1], Robert G. H. Bunce[2], Lisa R. Norton[1], Lindsay C. Maskell[1], David C. Howard[1], W. Andrew Scott[1], and Peter A. Henrys[1]

[1]Centre for Ecology and Hydrology, Lancaster Environment Centre, Bailrigg, Lancaster, LA1 4AP, UK
[2]Estonian University of Life Sciences, Kreuzwaldi 5, 51014 Tartu, Estonia

Correspondence to: C. M. Wood (clamw@ceh.ac.uk)

Received: 24 February 2017 – Discussion started: 6 March 2017
Revised: 13 June 2017 – Accepted: 15 June 2017 – Published: 20 July 2017

Abstract. The Countryside Survey (CS) of Great Britain provides a globally unique series of datasets, consisting of an extensive set of repeated ecological measurements at a national scale, covering a time span of 29 years. CS was first undertaken in 1978 to monitor ecological and land use change in Britain using standardised procedures for recording ecological data from representative 1 km squares throughout the country. The same sites, with some additional squares, were used for subsequent surveys of vegetation undertaken in 1990, 1998 and 2007, with the intention of future surveys. Other data records include soils, freshwater habitats and invertebrates, and land cover and landscape feature diversity and extents. These data have been recorded in the same locations on analogous dates. However, the present paper describes only the details of the vegetation surveys.

The survey design is a series of gridded, stratified, randomly selected 1 km squares taken as representative of classes derived from a statistical environmental classification of Britain. In the 1978 survey, 256 one-kilometre sample squares were recorded, increasing to 506 in 1990, 569 in 1998 and 591 in 2007. Initially each square contained up to 11 dispersed vegetation plots but additional plots were later placed in different features so that eventually up to 36 additional sampling plots were recorded, all of which can be relocated where possible (unless the plot has been lost, for example as a consequence of building work), providing a total of 16 992 plots by 2007. Plots are estimated to have a precise relocation accuracy of 85 %. A range of plots located in different land cover types and landscape features (for example, field boundaries) are included.

Although a range of analyses have already been carried out, with changes in the vegetation being related to a range of drivers at local and national scales, there is major potential for further analyses, for example in relation to climate change. Although the precise locations of the plots are restricted, largely for reasons of landowner confidentiality, sample sites are intended to be representative of larger areas, and many potential opportunities for further analyses remain.

Data from each of the survey years (1978, 1990, 1998, 2007) are available via the following DOIs: Countryside Survey 1978 vegetation plot data (https://doi.org/10.5285/67bbfabb-d981-4ced-b7e7-225205de9c96), Countryside Survey 1990 vegetation plot data (https://doi.org/10.5285/26e79792-5ffc-4116-9ac7-72193dd7f191), Countryside Survey 1998 vegetation plot data (https://doi.org/10.5285/07896bb2-7078-468c-b56d-fb8b41d47065), Countryside Survey 2007 vegetation plot data (https://doi.org/10.5285/57f97915-8ff1-473b-8c77-2564cbd747bc).

Published by Copernicus Publications.

1 Introduction

The Countryside Survey (CS) of Great Britain was initiated in the late 1970s to monitor ecological and land cover change using quantitative and repeatable methods. The history of the development of the methodology is given by Sheail and Bunce (2003). The survey is based on 1 km squares as a convenient sized unit, which had previously been tested in Cumbria (Bunce and Smith, 1978) and Shetland (Wood and Bunce, 2016) in the years preceding the first survey in 1978. The survey design is based on a series of dispersed, stratified, randomly selected 1 km squares from across Britain, which numbered 256 in 1978, 506 in 1990, 569 in 1998 and 591 in 2007. The stratification used was the statistical environmental classification of 1 km squares in Great Britain as described in Bunce et al. (1996b, c), and summarised in Sect. 2.2.

In the first survey, data were recorded from up to 11 vegetation plots of four different types, distributed through each of the squares (which form the main subject of the present paper), along with soil samples and land cover maps using standard classes which were later converted into standard habitat categories (Wood et al., 2012). Subsequent surveys including the vegetation component were undertaken in 1990, 1998 and most recently, 2007 (with an additional land use survey in 1984). During this period, additional vegetation plots have been placed in different land cover types and landscape features for policy objectives, eventually giving up to 36 more plots per square. Varying numbers of each vegetation plot type were initially placed in locations across each survey square according to rules outlined in Sect. 3. In subsequent surveys, these plots are repeated in these same fixed positions, except those such as on field margins, which are based on rules applied in the field. Details of the types of plot employed are described below, with an average of 29 plots being completed in each sample square. In addition to the vegetation plots described here, data are also recorded from linear features such as hedgerows, landscape elements such as veteran trees, areal broad habitats (Jackson, 2000) and related key species, soils and aquatic invertebrates (see Carey et al., 2008).

The survey as a whole provides a wealth of globally unique ecological data, consisting of an extensive range of measurements at a national scale, covering a time span of 29 years. From an international perspective, CS was a pioneer in surveys of its type. Although environmental surveys such as forest monitoring programmes were nothing new (e.g. United States Forest Service, http://www.fia.fs.fed.us/; Forest Survey of India, http://fsi.nic.in/), the integrated, systematic national monitoring of vegetation species, soils and landscape features across all land uses provided by CS was a novel concept, preceding programmes in many other countries particularly in Europe (e.g. Dramstad et al., 2002; Hintermann et al., 2002; Ståhl et al., 2011) and beyond (e.g. Burton et al., 2014).

2 Survey design: background, stratification and site selection

2.1 Background

As a result of the earlier work carried out in the 1970s on a regional scale (Wood and Bunce, 2016; Bunce and Smith, 1978), a sample unit of 1 km square was found to be an appropriate size for CS. It also forms the basic unit of the stratification framework described below. A 1 km square is small enough to survey in a reasonable period of time (1 week or less) and yet large enough to contain sufficient environmental features to allow differentiation between squares. Kilometre squares are used as a framework by the British national mapping agency, the Ordnance Survey, thus providing useful basemaps to aid surveyor navigation. A sampling unit of 1 km square is also widely used in other European projects (for example in Spain, Ortega et al., 2013), although in countries with small-scale landscapes, for example Northern Ireland, 0.25 km square has been adopted (Cooper et al., 2009).

2.2 Stratification

With over 240 000 of 1 km squares in Great Britain, a sampling approach was essential and a statistical environmental classification was constructed, from which the stratified, random samples of 1 km squares were taken. Due to the limitations of computing power and lack of readily available data at the time, the classification initially only covered a subset of 1212 one-kilometre squares, rather than the whole of Britain, based on a 15×15 km grid drawn from the National Grid defined by Britain's national mapping agency, the Ordnance Survey.

Altitude, climate, geology, human geography and location variables from each 1 km square were recorded manually for each 1 km square. Because the data were a mixture of variables (for example, altitude) and attributes (for example, geology) the variables were converted into four classes so that the database was suitable for analysis by indicator species analysis (ISA, now TWINSPAN (Hill and Šmilauer, 2005) and stopped at 32 classes. It is recognised that nowadays, with automated data capture, variables can be recorded for millions of 1 km squares, and recent environmental classifications (for example Metzger et al., 2008; Villoslada et al., 2016) have used principal component analysis and clustering. Jones and Bunce (1985) compared classifications of European climate using both methods and concluded that the results were comparable. Bunce et al. (2002) compared classifications for similar regions using different databases and analytical techniques and showed that the basic patterns were sufficiently close that policy makers would be able to have confidence in the results. The many multivariate techniques which are now available will give slightly different boundaries to classes but the core structure will always be identified. Finally, any inefficiencies in stratifications will be reflected in higher standard errors for the observed independent variables. The independent tests in papers such as Metzger et al. (2008) and Villoslada et al. (2016) are all highly significant and any improvements are likely to be marginal.

The resulting classes were described on the basis of average values of the environmental characteristics of the initial database, for example, altitude and rainfall (Bunce et al., 1996a, c).

A primary objective of the methodology is to reduce variation, as the classification divides the population into discrete strata which are then used to derive samples from which ecological parameters such as vegetation can be recorded. As a statistically robust method is used (i.e. ISA), it is possible to extrapolate the results from the sample sites into land class means, which can then be combined to describe an entire population (for example England, Scotland, Wales or Britain). The principles and development of this procedure are described by (Sheail and Bunce, 2003).

By 1990, all 1 km squares in Great Britain were classified into the same set of strata, which was not considered possible at the start of the 1970s. Known as the "Institute of Terrestrial Ecology (ITE) Land Classification of Great Britain" (Bunce et al., 1990, 1996a, b), it has developed over the 30-year period (Sheail and Bunce, 2003). The most recent modifications largely concern the incorporation of the requirement for country level reporting, separating Scotland (in 1998) and Wales (in 2007). The basic stratification still underpins the CS and the latest development of the original Land Classification consists of 45 classes (or strata), and is illustrated in Fig. 1, along with a map of the distribution of sampling locations (Fig. 2).

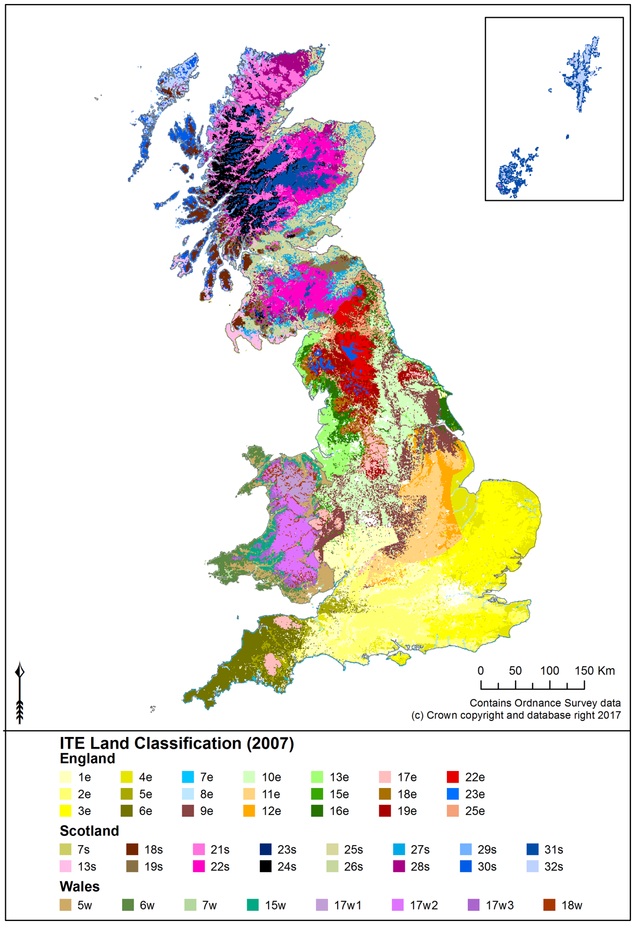


Figure 1. ITE Land Classification, 2007.

2.3 Site selection

Having constructed this initial stratification, the number of samples to be surveyed in the first (1978) survey was considered. Ideally, this number would depend on the size of the stratum (i.e. how many 1 km squares of the class occurred in Great Britain) and on the ecological variability within the stratum. Preliminary work had suggested that for ecological surveys of this type, at least eight samples per stratum were necessary in order to be representative of that stratum. Eight 1 km squares were therefore selected at random from each of the classes from the grid of classified squares. Thus the final sample for the first Great Britain survey was a gridded, stratified, random sample of 256 × 1 km squares. Surveying commenced in 1977, although the majority of squares were surveyed in 1978. Note that the location of the 1 km sample squares is not disclosed by agreement with land owners; the majority of the land in the sample squares is in private ownership. If the locations of the sites were made available, this would not only threaten future surveys but also prevent and future collaboration with the owners or their descendants. Furthermore, future land use decisions could be influenced and have an effect on the monitoring results. Thus this policy ensures that the squares do not attract additional research or land management activity that could potentially undermine their status as an unbiased, representative sample of the British countryside. In spite of the restrictions placed on the site locations, the data may be used in a wide range of analyses on a national or regional level, as described more fully in Sect. 7.

3 Sampling sites and plots

Initially, vegetation and soil data were recorded from five dispersed random ("X") plots in each 1 km square, which were located using a restricted randomisation procedure designed to reduce auto-correlation. Depending on the type of analyses in question, data users may wish to account for spatial auto-correlation, as in Baude et al (2016). However, in the majority of cases this is not an issue, as described by Betts et al. (2009), and as shown by in model checks in, for example, Henrys et al. (2015b). Vegetation was sampled from a further six plots placed along linear features (two hedgerow ("H") plots, two streamside ("S") plots, and two roadside ("R") plots). Plots have never been placed in built-up areas or be-

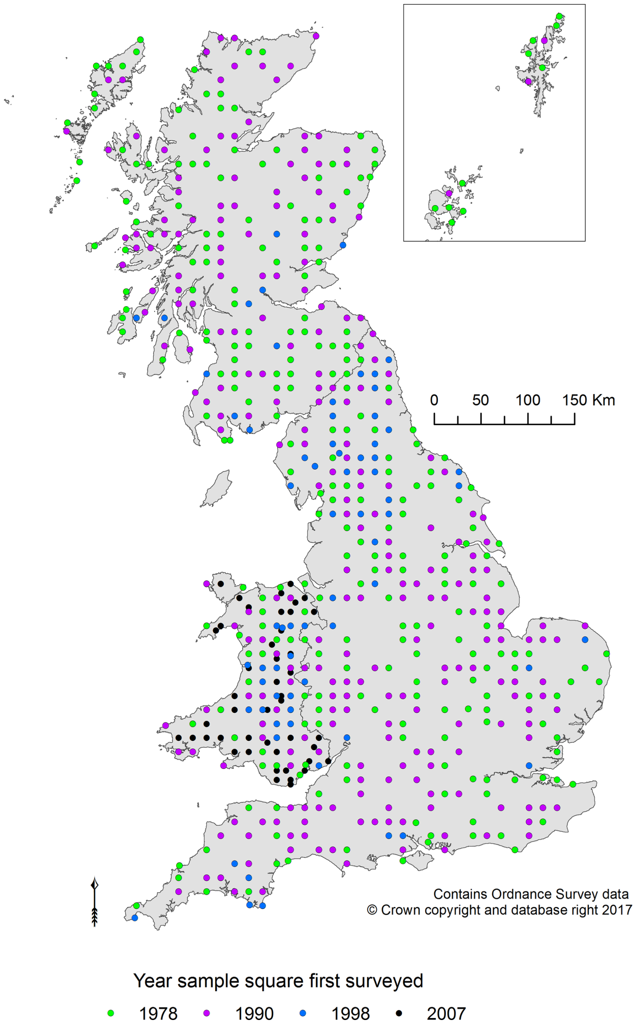


Figure 2. Map of Countryside Survey sampling locations across Britain.

low the mean high-water mark, and are only sited where the landowner has given permission. The types and total numbers of plots have increased over time from 1978 to 2007 along with the total number of CS 1 km squares surveyed. The total number of plots within squares varies depending on the landscape type and range of landscape features. Plots differ in size depending upon their type (Table 1). By 2007, the mean number of plots per square was 29 (min. 2, max. 47). The locations of all plots were mapped, together with measurements to local features, thus allowing them to be found again and re-recorded in the same place. Additional information ensuring the highest degree of accuracy when re-finding plots began to be recorded in the 1990 survey, as described in Sect. 4.2. The same plot locations have been repeated in all subsequent surveys (where appropriate), with additions.

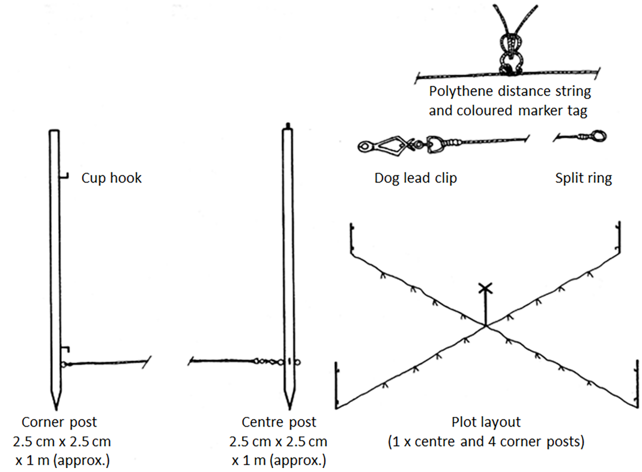


Figure 3. X plot construction.

4 Data collected

The vegetation survey involves recording plant species presence and abundance in different sizes and types of vegetation plot. In each vegetation plot, a complete list of all vascular plants and a selected range of readily identifiable bryophytes and macro-lichens is made (with the exception of D plots, which record woody species only). The field training course held before the surveys covered identification of difficult species, regular visits were made to survey teams by managers, and difficult specimens could be collected and sent to experts for identification. However, predetermined combinations of species may be recorded as aggregates reflecting known difficulties in their separation in the field (refer to Maskell et al., 2008; Barr, 1998). Cover estimates are made to the nearest 5 % for all species reaching at least an estimated 5 % cover. Presence is recorded if cover is less than 5 %. Canopy cover of overhanging trees and shrubs is also noted, even if individuals were not rooted within the plot. Additionally, general information about the plot is recorded to provide supporting information for analytical purposes as well as describing potential habitats such as glades and dead wood.

4.1 Plot types

X plots – large or main plots

The X plots are large nested plots designed to provide a random sample in proportion to the extent of the different vegetation types in each square and therefore in the wider countryside. X plots were pre-located before going into the field, with one plot being randomly placed into one of 5 equal sectors dividing the 1 km survey square. X plots typically sample the most common vegetation types. The X plot is 200 m^2 (14.14 × 14.14 m); the large size was adopted to obtain the maximum number of species within the plot. The methodol-

Table 1. Summary of vegetation plot types, sizes and numbers.

Code	Name	Other names	Where	Size	No. per square
Areal plots					
X[1]	Large	"Wally plot"; main	Random points in open polygons	200 m^2	5
Y[2+4]	Small	Targeted habitat	Uncommon vegetation types and in 2007, Priority habitats	4 m^2	Up to 5
U[3]	Unenclosed		Unenclosed broad habitats	4 m^2	Up to 10
Linear plots					
B[2]	Boundary		Adjacent to field boundaries and paired with X plots	10 × 1 m	5
A[3]	Arable		Arable field edges centred on each B plot	100 × 1 m	Up to 5
M[4+5]	Margin		Field margins	2 × 2 m	Up to 15
H[1]	Hedgerow		Alongside hedgerows	10 × 1 m	2
D[3]	Hedgerow diversity		Hedgerows/woody linear features	30 × 1 m	Up to 10
S[1]/W[2]	Streamside		Alongside water courses	10 × 1 m	5
R[1]/V[2]	Roadside		Alongside roads and tracks	10 × 1 m	5

[1] First recorded in 1978; [2] first recorded in 1990; [3] first recorded in 1998; [4] first recorded in 2007; [5] if there are five A plots in a square and wide margins.

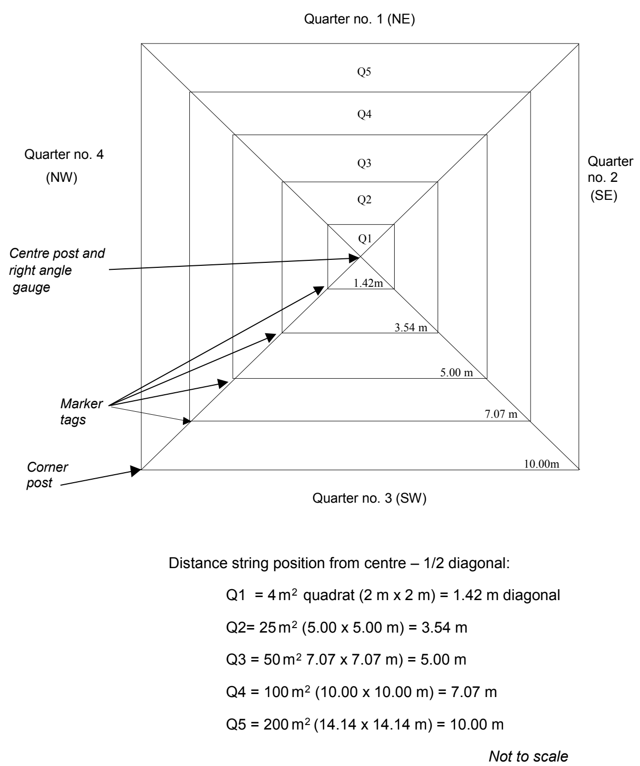


Figure 4. Layout of vegetation X plot.

ogy was originally produced for woodlands as described by Bunce and Shaw (1973) and was also used and found appropriate for strategic ecological survey, as described by Bunce and Smith (1978). The design of the plot not only aids a systematic search of the vegetation present but is straightforward to set out in the field, and ensures a standard area of the plot is covered on every occasion, making a square plot more advantageous than a circular plot in this case. The plot is set up by using a centre post and four corner posts, with a set of four strings tagged with markers at specified distances. The tagged strings form the diagonals of the square (as shown in Fig. 3). The diagonals should be orientated carefully at right angles and the plot should be orientated with the strings on the north/south, east/west axes. The different nested plots shown in Fig. 4 are marked by different coloured strings on the appropriate position of the diagonal. The design is to ensure that the whole plot is observed consistently and systematically, as unstructured search routines are more likely to lead to species being overlooked, as described as far back as 1940, by Hope-Simpson (1940).

Within the plot shown in Fig. 4, the first nest of the plot (2 × 2 m) is searched first. This procedure is then repeated for each nest of the quadrat, increasing the size each time and only recording additional species discovered in each larger nest (Fig. 4). In the final nest (the whole 200 m^2 plot), the percentage cover (to the nearest 5 %) of each species is also estimated. Estimates of cover for litter, wood, rock and bare ground are also included where present. In 2007, an additional 1 m^2 nest (not shown in Fig. 4) was introduced, in order to allow joint analysis of 1 m^2 plots being recorded in parallel as part of agri-environment scheme monitoring programs. This nest is located in the northernmost corner of the inner 4 m^2 nest (named nest "0"). Vegetation height, aspect and slope are also recorded. Soil samples are also taken at the same time, at the site of these plots; the procedure used for recording soil samples is given by Emmett et al. (2008) and is outside the remit of the present paper.

In arable fields where full access is not possible, for reasons of practicality species records are made from plots taken from an estimated 14 m square, starting 3 m into the crop, which avoids edge effects in most cases and minimises disturbance to the crop. Access is made using drill lines where possible in order to avoid trampling the crop. Overall cover is also estimated as in other land cover types. The relative uniformity of species within crops led to the adoption of this approach and the subsequent changes observed in species numbers in arable fields justified its use.

Y plots – small, targeted or habitat plots

These are small (2 × 2 m) plots located in less common vegetation types, usually of conservation interest, often occurring in small patches not sampled by other plot types. In 2007, additional Y plots also were placed in priority habitat (Maddock, 2008) patches that had also not been sampled by any other plot in the square. The Y plots are therefore important in sampling fragments of semi-natural habitat particularly in lowland landscapes where patches may be small and embedded in a matrix of intensive farmland. These plots are placed randomly by surveyors in suitable patches of vegetation (based on rules described in Maskell et al., 2008). Of all the plots recorded, they are most similar to the approach taken when positioning quadrats during National Vegetation Classification (NVC) (Rodwell, 2006) survey, where the location of the plot is designed to represent a vegetation unit perceived to be floristically distinct and homogenous. However, protocols for locating Y plots from 1998 onwards stipulated random location from within a larger extent of vegetation type in the 1 km square or from a number of patches representing the mapped land cover type. The validity of statistically analysing plots located with a degree of subjectivity is an ongoing matter of debate (see for example Lájer, 2007, and Palmer, 1993, for an illustration of analytical problems but also Ross et al., 2010, for counter-argument and examples of analysing temporal change in subjectively located samples).

U plots – unenclosed plots

These plots were introduced into the CS methodology for the first time in the 1998 survey to characterise the unenclosed broad habitats (Jackson, 2000) – these being calcareous and acid grasslands; bracken; dwarf shrub heath; fen, marsh and swamp; bog; montane; supra-littoral rock and sediment; and inland rock. Up to 10 plots were established in any unenclosed broad habitat types that occurred within the square (proportional to area), again placed randomly by surveyors. The plots are 2 × 2 m in all instances, regardless of the broad habitat in which they are located.

B plots – boundary plots

Boundary linear plots are recorded at a position on the boundary closest to each X plot and on a cardinal axis from it (i.e. north, south, east or west). A boundary is taken to be any linear physical feature that has a length greater than 20 m and which is an interface between the land cover of the 200 m^2 X plot and any other land cover type. This might include a hedge, wall, fence, ditch or embankment. These are linear 10 × 1 m plots.

A plots – arable field margin plots

Arable field margin plots were recorded for the first time in the 1998 survey. The purpose of establishing the plots was to record the arable weed population at the edge of cultivated fields and any subsequent changes. Theses plots relate only to the edge of fields and are quite distinct from the (arable) X (main) plots which are actually in the crop. They contribute an important source of biodiversity not present in the arable main plots, which cover the overall composition of arable crops because as described above, the margin is specifically excluded. The uptake of "conservation headland" options for arable field management under agri-environment schemes may further enhance species diversity in A plots. The plots are 100 m long by 1 m wide and located adjacent to B plots which border arable fields, up to a maximum of five per square. They always sample the first 1 m of cultivated land moving away from the perennial-dominated margin.

M plots – margin plots

M plots are small (2 × 2 m) square plots and were new in the 2007 survey. They are associated with B plots where an A plot is present, and the number depends on the widths of the margins present, with up to three per field. They are designed to record the quality of new arable field margins that form part of the agri-environment agreements on farms and other margins put in without agri-environment support. These margins are additional to the cross-compliance margin (not relevant in Wales), which is a 2 m margin measured from the centre of the hedge. The most common types of margin likely to be encountered are perennial grass margins, with or without supplementary wildflowers. Other rarer types include, uncropped strips, usually cultivated each year (regenerating from the seedbank); wild bird seed cover, e.g. kale, quinoa; and pollen and nectar mixes, usually with a high proportion of legumes.

H plots – hedgerow plots

H plots are linear 10 × 1 m plots running alongside managed woody linear features ("WLFs", hedgerows). Within H plots, species associated with the managed WLFs are recorded. When first recorded in 1978, the plot positions were located as close as possible to the two X plots in each square which were furthest apart.

D plots – hedgerow diversity plots

Hedgerow diversity plots were recorded for the first time in 1998. The overall purpose was to set up a baseline of plots to monitor woody species diversity in WLFs. One D plot is placed on each WLF in the square, up to a maximum of 10 plots. As well as providing information on woody species diversity, the data collected in D plots also help to provide an assessment of the condition of hedgerows and other WLFs by providing vital information about the size of the WLFs, gappiness, levels of disturbance and species composition. Each plot is 30 m long and includes the full width of the WLF.

S/W plots – streamside plots

The term “streamside plot” denotes linear plots which lie alongside running water features (mainly rivers and streams, but also canals and ditches). The S and W prefixes refer to the different origins of the plots: two Streamside (S) plots were established in 256 of the 1 km squares in 1978, located as close as possible to the two X plots in each square which were furthest apart. W plots were up to three additional waterside plots, placed in all squares in 1990 to increase representation of other waterside types. These are linear 10 × 1 m plots.

R/V plots – roadside and verge plots

The term “roadside plot” denotes those linear plots which lie alongside transport routes (mainly roads and tracks). The R and V prefixes refer to the different origins of the plots: two roadside (R) plots were established in 256 of the 1 km squares in 1978, located as close as possible to the two X plots in each square which were furthest apart. V plots are up to three additional verge plots first placed in 1 km squares in 1990 to increase representation of other transport types. These are linear 10 × 1 m plots.

4.2 Plot relocation

To analyse change, it is important to relocate the exact same sampling plot locations in successive surveys. The data from the repeated vegetation plots provide a globally unique dataset allowing large-scale yet fine-grained change in overall vegetation and the state or condition of the broad and priority habitats over time to be documented at four points over the last 29 years. There are no other national data sets that cover entire landscapes, including constituent habitats over such a long period of time. In practice, there are actually very few long-term studies of vegetation change. Those existing are usually either opportunistic, because some local recording has given a precise location, for example Dunnett et al. (1998) on a roadside verge in Bibury in Gloucestershire, or pertain to specific habitats, such as the Park Grass Experiment at Rothamsted (Silvertown et al., 2006).

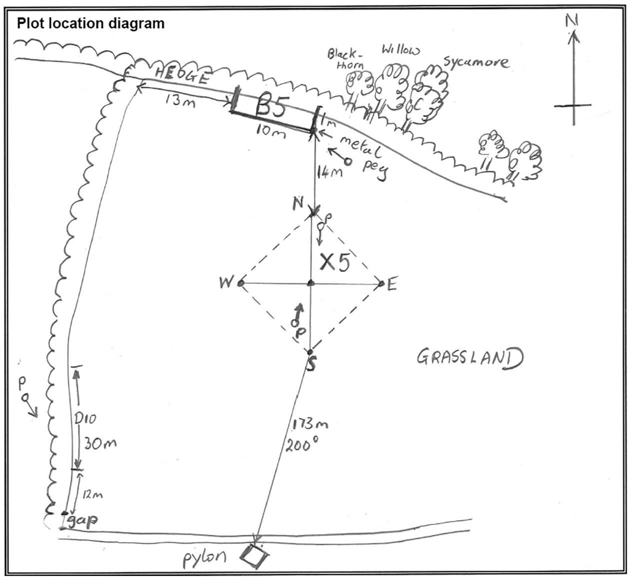


Figure 5. Example of a plot sketch map.

During the surveys, plot locations have been recorded on paper using a sketch map with measurements from distinguishing landscape features (Fig. 5), and by taking at least two photographs (see Fig. 6 for an example), preferably also including key landscape features in proximity to the plot. In addition to these, permanent metal plates or wooden stakes were introduced in the 1990 survey. In 1998, a GPS position was recorded in some remote squares, which assisted locating plots again in 2007. In 2007, the plot locations were recorded via the ruggedised field computers using the in-built GPS (where a GPS signal was available). Surveyors are also able to record whether the plots have been re-found adequately or otherwise. Circumstances where a plot may not be repeated might include an area becoming built-up, a feature having been removed or a land owner refusing access to the land containing the plot. Using a combination of metal plates, photographs and sketch maps, plots are estimated to have a precise relocation accuracy of 85–86 % (Prosser and Wallace, 2008). (See Prosser and Wallace, 2008, for further analysis regarding this issue.)

5 Data quality

Each field survey was carried out by teams of experienced botanical surveyors, and was preceded by an intensive training course, ensuring high standards and consistency of methodology, effort, identification and recording across CS according to criteria laid out in the field handbooks (Maskell et al., 2008; Barr, 1998, 1990; Bunce, 1978). During the surveys, survey teams were initially supervised and later mon-

Figure 6. Example of a plot photograph.

itored by experienced project staff in order to control data quality.

Data were recorded on waterproof paper sheets in 1978, 1990 and 1998 and were consequently transferred from the original field sheets to spreadsheets, using a "double-punch" method to minimise errors in data entry. They were checked using range and format checks, and corrected to produce a final validated copy. In 2007, a new electronic data capture method was developed by the Centre for Ecology & Hydrology and used in CS for the first time. The move to electronic methods created greater efficiency in terms of data entry and also eliminated a potentially significant source of error. Improvements to data quality also resulted from the inclusion of mandatory data entry fields for each plot.

In terms of assessing the actual level of botanical expertise in the field surveys, quality assurance (QA) reports were completed by independent botanists for the surveys in 1990, 1998 and 2007 (Prosser and Wallace, 2008, 1999, 1992). These reports have been a vital tool in assessing and validating the quality of the botanical record in each CS. Paired species records from a subset of plots (the QA plots) have been analysed in a number of ways to measure the consistency of recording effort within each survey. In all three surveys the QA assessors found more species than the CS field teams, yet in both the 1990 and 1998 assessments, the results showed that there was no bias in the species composition of the vegetation recorded, as described by Prosser and Wallace (2008). In 2007, the QA analysis appeared to show a decline in the quality of botanical recording. However, this was possibly due to less comprehensive recording of common bryophytes than in previous surveys, but subsequent analyses determined that the bias was not significant (Scott et al., 2008; Smart et al., 2008). Users of the data could remove bryophytes from analyses if they were concerned by this feature of the database. Errors attributable to use of the electronic data capture software were minor and not significant (Prosser and Wallace, 2008).

6 Methodological development

The now established method of CS, using a stratified random series of samples, was developed over two decades by what was then the Institute of Terrestrial Ecology as described by Sheail and Bunce (2003). The first national series of stratified random samples was the 1971 Woodland Survey (Wood et al., 2015) and strategic sampling at the landscape level was subsequently used successfully in defining the range of variation in vegetation in regional surveys in Cumbria and Shetland (Bunce and Smith, 1978; Wood and Bunce, 2016). These methods have now been proven as a successful national vegetation monitoring strategy incorporating four surveys across nearly 30 years. Minor modifications to the methods have more recently been used for a comprehensive ecological survey of Wales (2013–), the Glastir Monitoring and Evaluation Programme (GMEP) (Emmett and GMEP team, 2014).

Since the first survey in 1978, the methods have gradually developed to incorporate contemporary technologies, for example, the introduction of GPS in 1998, and the use of ruggedised field computers with internal GPS to record the location and species composition of the vegetation plots in 2007. Over time, the development of geographical information systems (GISs) has greatly facilitated both the efficiency of storage of ecological spatial data, and also the types of analyses that can be undertaken. It is now possible to perform much wider analyses than previously, using a range of ancillary explanatory datasets, as described in the Integrated Assessment Report for the Countryside Survey (Smart et al., 2010a). The underlying principles of the Countryside Survey methodology provide an ideal framework for the planning of large-scale monitoring, not only in Britain but across Europe and worldwide, as discussed in Wood and Bunce (2016).

It has now been a decade since the last survey, and current funding constraints mean that the traditional cycle of large one-off national surveys taking place roughly 1 year in every decade is likely to need revising. Various options are available for repeating all or parts of the survey. A rolling program over several years is attractive because it spreads the financial load. It also allows inter-annual effects of differences in the weather and variation in recorder effort to be more robustly estimated and separated from long-term trends. A Markov chain approach could be used to examine possible outcomes from the time series of plots (for example, Balzter, 2000). Between 2013 and 2016, CS methods have already been applied in an annual rolling program to monitor the effects of the Glastir agri-environment scheme across Wales (https://gmep.wales/).

Plot numbers could be rationalised according to the desired results. Using previous data, it is possible to identify the optimal numbers of plots required by plot type, vegetation type and region in order to provide data on specific criteria, for example, species richness change at Great Britain level by plot type. Less costly options for maximising the use of the existing surveys in future surveillance have been suggested as part of the Future Options review for national monitoring in Wales (Emmett et al., 2016a). However, the feasibility of these options has yet to be determined.

7 Use of the data

The Countryside Survey provides a unique and well-utilised resource, offering potential for a wide range of analyses at different temporal and spatial scales. A major benefit of the programme is the co-registration of a wide range of recorded ecological variables (i.e. soil, vegetation, land use, freshwater). In parallel to its direct policy application, a vibrant and productive research agenda has used CS vegetation data often in combination with other datasets to produce improved understanding about the significance and causes of large-scale but finely resolved ecological change in Britain. Questions can be broadly categorised as "What has changed and where?", "What are the drivers of change?", "Is the change important?" and "Can we use forecast future change?".

As the data from the vegetation plots are intended to be representative of the larger areas in which they are located (i.e. land class), the restrictions on the precise locations of the plots need not restrict potential analyses. For example, the design of the survey is such that the data are intended to be extrapolated to a land class and, ultimately, a national level. Data may also be used in conjunction with other co-located variables (for example, soils data; Emmett et al., 2016c) to examine inter-relationships.

7.1 Key findings

Key findings and fundamental questions about the extent and condition of terrestrial broad/priority habitats are addressed in the reporting round to policy makers that has followed each survey (e.g. Haines-Young et al., 2000; Carey et al., 2008; Smart et al., 2009; Norton et al., 2009; Emmett et al., 2010; Bunce, 1979; Barr et al., 1993) and elsewhere (Smart et al., 2003; Norton, 2012). Overall changes in plant species richness formed part of a trend in species loss (8 %) across Great Britain between 1978 and 2007 (although this measure is a simple one, it is readily understood and appreciated by policy makers; however, it does need to be supported by the more detailed ecological analyses described in Sect. 7.2). Woody species increased in vegetation associated with landscape boundaries by 14 % between 1998 and 2007 and by nearly 80 % in Great Britain between 1978 and 2007.

The most commonly recorded species in CS, ryegrass (*Lolium perenne*), was the same in 2007 as in both 1998 and 1990. The top 10 most commonly recorded species in 2007 also included stinging nettle (*Urtica dioica*), hawthorn (*Crataegus monogyna*), and bramble (*Rubus fruticosus*) all of which increased between 1998 and 2007.

Long-term change in vegetation from 1978 to 2007 has also been assessed using a range of condition measures (Table 2). In open countryside in Great Britain, between 1998 and 2007 plant species that prefer wetter conditions increased, while those preferring fertile soils and high pH decreased. In the period 1978 to 2007, an increase in species preferring wetter conditions was the most consistent signal in plots sampling different parts of the landscape across all countries.

7.2 Wider uses of data to date

After CS in 2007, the data continued to have a substantial impact, contributing to many areas of the UK National Ecosystem Assessment (NEA) (Watson et al., 2011), which articulated ecological status and change in terms of ecosystem services (ESs). This was the first analysis of the UK's natural environment in terms of the benefits it provides to society and continuing economic prosperity. Soils, vegetation, headwater stream and land cover data from Countryside Survey were also jointly analysed with a range of explanatory variable datasets to produce new indicators and analysis of potential ES delivery in the Integrated Assessment project that marked the final phase of reporting after the 2007 survey (Smart et al., 2010a; Norton et al., 2012).

Subsequently, CS plot data have been used in conjunction with land cover map data (Morton et al., 2011) and wider environmental datasets as part of a natural capital mapping tool which has been used, alongside other modelling techniques, to produce maps of natural capital for policy makers (https://eip.ceh.ac.uk/naturalengland-ncmaps) and to help in understanding the factors influencing spatial differences in ES delivery (Henrys et al., 2015a; Norton et al., 2016). Analysis demonstrated fundamental trade-offs between ecosystem productivity and soil carbon concentration while a range of biodiversity indicators appeared to peak at intermediate levels of productivity (Maskell et al., 2013). The novel inclusion of dynamic ecosystem model estimates of productivity provided both the foundation and research direction for ongoing work that has sought to develop dynamic models of natural capital and ES delivery (Emmett et al., 2016b; Smart et al., 2017; Rowe et al., 2016).

CS datasets have also made a unique contribution to the development of plant species niche models for ecosystem dominants and many rare species in Britain (Hill et al., 2017; Henrys et al., 2015b, c; Smart et al., 2010b, c). The policy motivation for this originally was detection and modelling of the effects of atmospheric pollutant deposition (De Vries et al., 2010; Stevens et al., 2016).

The statistically robust, national scale of the CS vegetation data makes it ideally placed to detect realistically scaled relationships between global change drivers, such as pollutant deposition (for example van den Berg et al., 2016; Maskell et

Table 2. Change in the characteristics of all types of vegetation in 200 m^2 main plots in Great Britain between 1978 and 2007. Arrows denote significant change ($p < 0.05$) in the direction shown.

	Mean values				Direction of significant changes 1978–2007
Vegetation condition measures	1978	1990	1998	2007	GB
Species richness (no. of species)	17.1	16.5	16.2	15.7	↓
Light score	6.98	6.95	6.95	6.95	
Fertility score	4.53	4.64	4.61	4.55	
Ellenberg pH score*	5.07	5.17	5.14	5.09	
Moisture score	5.75	5.71	5.77	5.82	↑

* Ellenberg (1988)

al., 2010; Smart et al., 2005a; Stevens et al., 2009) as well as other drivers of eutrophication and land management change (Smart et al., 2012, 2002, 2003, 2005b, 2006a, b). While research into the causes and consequences of eutrophication was a response to clear policy interest, analysis of CS vegetation data has also contributed evidence in response to concerns over the causes and consequences of loss of pollinators in north-west Europe and Britain (Smart et al., 2000; Carvell et al., 2006; Baude et al., 2016).

Habitat specific studies, such as those relating to woodlands (for example Petit et al., 2004; Kimberley et al., 2013, 2016) and hedgerows, McCollin et al., 2000; Garbutt and Sparks, 2002; Critchley et al., 2013) have been facilitated through the use of CS data. Interesting conclusions have been made through use of the data with regard to increasing numbers of non-native invasive species (Chytrý et al., 2008; Maskell et al., 2006).

The results and information derived from CS can often be set into wider contexts, for example, European (Chytrý et al., 2008; Metzger et al., 2013), or in relation to other monitored datasets (Rose et al., 2016; Scott et al., 2010; Carey et al., 2002; Rhodes et al., 2015).

Data from the 1990 survey were used in the development of a statistically based British vegetation classification, termed the Countryside Vegetation Classification (CVS) as described in Bunce et al. (1999). This led to the development of a computer system termed MAVIS (Modular Analysis of Vegetation Information System), enabling classification of any lists of species from plots into the CVS but also into the phytosociological classes of the National Vegetation Classification (Rodwell, 2006). The software is publicly available (Smart and DART Computing, 2017).

8 Data availability

The datasets have been assigned digital object identifiers and users of the data must reference the data as follows:

– Barr, C. J., Bunce, R. G. H., Smart, S. M., and Whittaker, H. A.: Countryside Survey 1978 vegetation plot data, NERC Environmental Information Data Centre, https://doi.org/10.5285/67bbfabb-d981-4ced-b7e7-225205de9c96, 2014.

– Barr, C. J., Bunce, R. G. H., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Smart, S. M., Stuart, R. C., and Wood, C. M.: Countryside Survey 1990 vegetation plot data, NERC Environmental Information Data Centre, https://doi.org/10.5285/26e79792-5ffc-4116-9ac7-72193dd7f191, 2014.

– Barr, C. J., Bunce, R. G. H., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Shield, E. R., Smart, S. M., Stuart, R. C., Watkins, J. W., and Wood, C. M.: Countryside Survey 1998 vegetation plot data, NERC Environmental Information Data Centre, https://doi.org/10.5285/07896bb2-7078-468c-b56d-fb8b41d47065, 2014.

– Bunce, R. G. H., Carey, P. D., Maskell, L. C., Norton, L. R., Scott, R. J., Smart, S. M., and Wood, C. M.: Countryside Survey 2007 vegetation plot data, NERC Environmental Information Data Centre, https://doi.org/10.5285/57f97915-8ff1-473b-8c77-2564cbd747bc, 2014.

The datasets are available from the CEH Environmental Information Data Centre Catalogue (https://eip.ceh.ac.uk/data). Datasets are provided under the terms of the Open Government Licence (http://eidchub.ceh.ac.uk/administration-folder/tools/ceh-standard-licence-texts/ceh-open-government-licence/plain, http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/). The metadata are stored in the ISO 19115 (2003) schema (International Organization for Standardization, 2015) in the UK Gemini 2.1 profile (UK GEMINI, http://www.agi.org.uk/join-us/agi-groups/standards-committee/uk-gemini). Users of the datasets will find the following documents useful (supplied as supporting documentation with the datasets): the Sampling Strategy for Countryside Survey (Barr and Wood,

2011) and the field survey handbooks (Barr, 1990, 1998; Bunce, 1978; Maskell et al., 2008).

9 Conclusions

The vegetation data recorded during the Countryside Survey of Great Britain are an invaluable national resource, which, over the years, has been exploited in a large number of ways. The data are collected in a statistically robust and quality controlled manner, follow standard, repeatable methods and cover wide temporal and spatial scales. As consequence of this, the data present a unique opportunity for inclusion in a wide range of analyses and models. The intention is that a repeat survey will be undertaken in the near future (indeed a sub-sample of plots (the majority being located in Wales) have already been surveyed in the summer of 2016, largely as part of the Glastir Monitoring and Evaluation Programme, Emmett and GMEP team, 2014). As a decade has now passed since the most recent full survey, an addition to this long-term national resource is becoming increasingly timely, particularly in these current times of political, socio-economic and climatic change.

Author contributions. CMW prepared the manuscript with significant contributions from all co-authors, and is the current database manager for the Land Use Research Group at CEH Lancaster. RGHB designed the sampling framework and survey strategy in 1978. RGHB, SMS, LCM, LRN and DCH have all been part of the Countryside Survey co-ordination team for at least one survey, with WAS and PAH contributing statistical support.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank all the land owners who kindly gave permission to survey their holdings in the survey sample squares in 1978, 1990, 1998 and 2007. Without their co-operation and assistance the Countryside Survey would not exist. We also acknowledge and thank all the field surveyors involved in each field campaign. We are grateful to two anonymous reviewers and the editor, Falk Huettmann, whose comments improved the manuscript considerably.

The most recent Countryside Survey was funded largely by the Natural Environment Research Council, the Department for Environment Food and Rural Affairs and the Scottish and Welsh Governments.

Edited by: Falk Huettmann
Reviewed by: two anonymous referees

References

Balzter, H.: Markov chain models for vegetation dynamics, Ecol. Model., 126, 139–154, https://doi.org/10.1016/S0304-3800(00)00262-3, 2000.

Barr, C.: Countryside Survey 1990 field handbook, Institute of Terrestrial Ecology, Grange-over-Sands, 1990.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Fuller, R. M., Furse, M. T., Gillespie, M. K., Groom, G. B., Hallam, C. J., Hornung, M., Howard, D. C., and Ness, M. J.: Countryside Survey 1990: main report, in: Countryside 1990, Vol. 2, Department of the Environment, London, 1993.

Barr, C. J.: Countryside Survey 2000 Field Handbook, Institute of Terrestrial Ecology, Grange-over-Sands, 1998.

Barr, C. J. and Wood, C. M.: The Sampling Strategy for Countryside Survey (up to 2007), 2nd Edn., Detr Contract No. CR0212, Lancaster, 2011.

Barr, C. J., Bunce, R. G. H., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Smart, S. M., Stuart, R. C., and Wood, C. M.: Countryside Survey 1990 vegetation plot data, NERC Environmental Information Data Centre, https://doi.org/10.5285/26e79792-5ffc-4116-9ac7-72193dd7f191, 2014a.

Barr, C. J., Bunce, R. G. H., Gillespie, M. K., Howard, D. C., Maskell, L. C., Norton, L. R., Scott, R. J., Shield, E. R., Smart, S. M., Stuart, R. C., Watkins, J. W., and Wood, C. M.: Countryside Survey 1998 vegetation plot data, NERC Environmental Information Data Centre, https://doi.org/10.5285/07896bb2-7078-468c-b56d-fb8b41d47065, 2014b.

Barr, C. J., Bunce, R. G. H., Smart, S. M., and Whittaker, H. A.: Countryside Survey 1978 vegetation plot data, NERC Environmental Information Data Centre, https://doi.org/10.5285/67bbfabb-d981-4ced-b7e7-225205de9c96, 2014c.

Baude, M., Kunin, W. E., Boatman, N. D., Conyers, S., Davies, N., Gillespie, M. A., Morton, R. D., Smart, S. M., and Memmott, J.: Historical nectar assessment reveals the fall and rise of floral resources in Britain, Nature, 530, 85–88, https://doi.org/10.1038/nature16532, 2016.

Betts, M. G., Ganio, L. M., Huso, M. M. P., Som, N. A., Huettmann, F., Bowman, J., and Wintle, B. A.: Comment on "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review", Ecography, 32, 374–378, https://doi.org/10.1111/j.1600-0587.2008.05562.x, 2009.

Bunce, R., Barr, C., Gillespie, M., Howard, D., Scott, W., Smart, S., van de Poll, H., and Watkins, J.: ECOFACT 1 – Vegetation of the British Countryside, HMSO, London, 1999.

Bunce, R. G. H.: UK Ecological Survey. Handbook of Field Methods, Insitute of Terrestrial Ecology, Grange-over-Sands, 1978.

Bunce, R. G. H.: Ecological survey of Britain, Institute of Terrestrial Ecology, Cambridge, 74–75, 1979.

Bunce, R. G. H. and Shaw, M. W.: National survey of semi-natural woodlands, in: Merlewood Research Station Report for 1970–72, The Nature Conservancy, Grange-over-Sands, 1973.

Bunce, R. G. H. and Smith, R. S.: An ecological survey of Cumbria, Structure Plan Working Paper, 4, Cumbria County Council and Lake District Special Planning Board, Kendal 1978.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE land classification of Great

Britain 1990, NERC Environmental Information Data Centre, https://doi.org/10.5285/ab320e08-faf5-48e1-9ec9-77a213d2907f, 1990.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE Merlewood Land Classification of Great Britain, J. Biogeogr., 23, 625-634, https://doi.org/10.1111/j.1365-2699.1996.tb00023.x, 1996a.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: Land classification for strategic ecological survey, J. Environ. Manage., 47, 37–60, https://doi.org/10.1006/jema.1996.0034, 1996b.

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., and Howard, D. C.: The ITE Land Classification: providing an environmental stratification of Great Britain, Environ. Monit. Assess., 39, 39–46, https://doi.org/10.1007/978-94-009-1653-1_5, 1996c.

Bunce, R. G. H., Carey, P. D., Elena-Rossello, R., Orr, J., Watkins, J., and Fuller, R.: A comparison of different biogeographical classifications of Europe, Great Britain and Spain, J. Environ. Manage., 65, 121–134, https://doi.org/10.1006/jema.2002.0533, 2002.

Bunce, R. G. H., Carey, P. D., Maskell, L. C., Norton, L. R., Scott, R. J., Smart, S. M., and Wood, C. M.: Countryside Survey 2007 vegetation plot data, NERC Environmental Information Data Centre, https://doi.org/10.5285/57f97915-8ff1-473b-8c77-2564cbd747bc, 2014.

Burton, A. C., Huggard, D., Bayne, E., Schieck, J., Sólymos, P., Muhly, T., Farr, D., and Boutin, S.: A framework for adaptive monitoring of the cumulative effects of human footprint on biodiversity, Environ. Monit. Assess., 186, 3605–3617, 2014.

Carey, P., Barnett, C., Greenslade, P., Hulmes, S., Garbutt, R., Warman, E., Myhill, D., Scott, R., Smart, S., and Manchester, S.: A comparison of the ecological quality of land between an English agri-environment scheme and the countryside as a whole, Biol. Conserv., 108, 183–197, https://doi.org/10.1016/S0006-3207(02)00105-2, 2002.

Carey, P. D., Wallis, S., Chamberlain, P. M., Cooper, A., Emmett, B. A., Maskell, L. C., McCann, T., Murphy, J., Norton, L. R., Reynolds, B., Scott, W. A., Simpson, I. C., Smart, S. M., and Ullyett, J. M.: Countryside Survey: UK Results from 2007, NERC/Centre for Ecology & Hydrology, Lancaster, 2008.

Carvell, C., Roy, D. B., Smart, S. M., Pywell, R. F., Preston, C. D., and Goulson, D.: Declines in forage availability for bumblebees at a national scale, Biol. Conserv., 132, 481–489, https://doi.org/10.1016/j.biocon.2006.05.008, 2006.

Chytrý, M., Maskell, L. C., Pino, J., Pyšek, P., Vilà, M., Font, X., and Smart, S. M.: Habitat invasions by alien plants: a quantitative comparison among Mediterranean, subcontinental and oceanic regions of Europe, J. Appl. Ecol., 45, 448–458, https://doi.org/10.1111/j.1365-2664.2007.01398.x, 2008.

Cooper, A., McCann, T., and Rogers, D.: Northern Ireland countryside survey 2007: broad habitat change 1998–2007, Northern Ireland Environment Agency, Coleraine, 2009.

Critchley, C., Wilson, L., Mole, A., Norton, L., and Smart, S.: A functional classification of herbaceous hedgerow vegetation for setting restoration objectives, Biodivers. Conserv., 22, 701–717, https://doi.org/10.1007/s10531-013-0440-5, 2013.

De Vries, W., Wamelink, G., Van Dobben, H., Kros, J., Reinds, G., Mol-Dijkstra, J., Smart, S., Evans, C., Rowe, E., and Belyazid, S.: Use of dynamic soil-vegetation models to assess impacts of nitrogen deposition on plant species composition: an overview, Ecol. Appl., 20, 60–79, https://doi.org/10.1890/08-1019.1, 2010.

Dramstad, W. E., Fjellstad, W., Strand, G.-H., Mathiesen, H. F., Engan, G., and Stokland, J. N.: Development and implementation of the Norwegian monitoring programme for agricultural landscapes, J. Environ. Manage., 64, 49–63, 2002.

Dunnett, N., Willis, A., Hunt, R., and Grime, J.: A 38-year study of relations between weather and vegetation dynamics in road verges near Bibury, Gloucestershire, J. Ecol., 86, 610–623, https://doi.org/10.1046/j.1365-2745.1998.00297.x, 1998.

Ellenberg, H.: Vegetation ecology of central Europe, Cambridge University Press, Cambridge, 1988.

Emmett, B., Frogbrook, Z., Chamberlain, P., Griffiths, R., Pickup, R., Poskitt, J., Reynolds, B., Rowe, E., Spurgeon, D., and Rowland, P.: CS Technical Report No. 3/07: Soils Manual, Centre for Ecology & Hydrology Wallingford, UK, 2008.

Emmett, B. A., Reynolds, B., Chamberlain, P. M., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A. J., Poskitt, J., Potter, E., Robinson, D. A., Scott, A., Wood, C., and Woods, C.: Countryside Survey: Soils Report from 2007, Centre for Ecology & Hydrology, Wallingford, UK, 2010.

Emmett, B. E. and GMEP team: Glastir Monitoring & Evaluation Programme, First Year Annual Report to Welsh Government (Contract reference: C147/2010/11), NERC/Centre for Ecology & Hydrology, Bangor, 442 pp., 2014.

Emmett, B., Bell, C., Chadwick, D., Cheffings, C., Henrys, P., Prosser, H., Siriwardena, G., Smart, S., and Williams, B.: Options for a new integrated natural resources monitoring framework for Wales, Phase 1 project report, NERC/Centre for Ecology & Hydrology, Bangor, UK, 2016a.

Emmett, B., Cooper, D., Smart, S., Jackson, B., Thomas, A., Cosby, B., Evans, C., Glanville, H., McDonald, J. E., and Malham, S. K.: Spatial patterns and environmental constraints on ecosystem services at a catchment scale, Sci. Total Environ., 572, 1586–1600, https://doi.org/10.1016/j.scitotenv.2016.04.004, 2016b.

Emmett, B., Reynolds, B., Chamberlain, P. M., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A. J., Poskitt, J., Potter, E., Robinson, D. A., Scott, A., Wood, C. M., Woods, C.: Soil physico-chemical properties 2007, Countryside Survey, NERC Environmental Information Data Centre, available at: https://doi.org/10.5285/79669141-cde5-49f0-b24d-f3c6a1a52db8, 2016c.

Garbutt, R. and Sparks, T.: Changes in the botanical diversity of a species rich ancient hedgerow between two surveys (1971–1998), Biol.Conserv., 106, 273–278, https://doi.org/10.1016/S0006-3207(01)00253-1, 2002.

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

Henrys, P., Bee, E., Watkins, J., Smith, N., and Griffiths, R.: Mapping natural capital: optimising the use of national scale datasets, Ecography, 38, 632–638, https://doi.org/10.1111/ecog.00402, 2015a.

Henrys, P., Smart, S., Rowe, E., Jarvis, S., Fang, Z., Evans, C., Emmett, B., and Butler, A.: Niche models for British plants and lichens obtained using an ensemble approach, New J. Botany, 5, 89–100, 2015b.

Henrys, P. A., Butler, A., Jarvis, S., Smart, S. M., and Fang, Z.: MultiMOVE Model: Ensemble niche modelling of British vegetation v2.0.1, NERC Environmental Information Data Centre, https://doi.org/10.5285/94ae1a5a-2a28-4315-8d4b-35ae964fc3b9, 2015c.

Hill, L., Hector, A., Hemery, G., Smart, S., Tanadini, M., and Brown, N.: Abundance distributions for tree species in Great Britain: a two-stage approach to modeling abundance using species distribution modeling and random forest, Ecol. Evol., 7, 1043–1056, https://doi.org/10.1002/ece3.2661, 2017.

Hill, M. O. and Šmilauer, P.: TWINSPAN for Windows version 2.3., Centre for Ecology and Hydrology, University of South Bohemia, Wallingford, 2005.

Hintermann, U., Weber, D., Zangger, A., and Schmill, J.: Biodiversity Monitoring in Switzerland, BDM – Interim Report, Berne, 2002.

Hope-Simpson, J.: On the errors in the ordinary use of subjective frequency estimations in grassland, J. Ecol., 28, 193–209, https://doi.org/10.2307/2256169, 1940.

International Organization for Standardization, ISO 19115:2003: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020 (last access: 12 October 2015), 2015.

Jackson, D.: Guidance on the Interpretation of the Biodiversity Broad Habitat Classification (Terrestrial and Freshwater Types): Definitions and the Relationship with Other Habitat Classifications (JNCC Report, No 307), JNCC, available at: http://jncc.defra.gov.uk/page-2433, 73 pp., 2000.

Jones, H. E. and Bunce, R. G. H.: A preliminary classification of the climate of Europe from temperature and precipitation records, J. Environ. Manage., 20, 17–29, 1985.

Kimberley, A., Blackburn, G. A., Whyatt, J. D., Kirby, K., and Smart, S. M.: Identifying the trait syndromes of conservation indicator species: how distinct are British ancient woodland indicator plants from other woodland species?, Appl. Veg. Sci., 16, 667–675, https://doi.org/10.1111/avsc.12047, 2013.

Kimberley, A., Blackburn, G. A., Whyatt, J. D., and Smart, S. M.: How well is current plant trait composition predicted by modern and historical forest spatial configuration?, Ecography, 39, 67–76, https://doi.org/10.1111/ecog.01607, 2016.

Lájer, K.: Statistical tests as inappropriate tools for data analysis performed on non-random samples of plant communities, Folia Geobotanica, 42, 115–122, https://doi.org/10.1007/BF02893878, 2007.

Maddock, A.: UK Biodiversity Action Plan; Priority Habitat Descriptions, available at: http://jncc.defra.gov.uk/PDF/UKBAP_PriorityHabitatDesc-Rev2011.pdf, 2008.

Maskell, L., Firbank, L., Thompson, K., Bullock, J., and Smart, S.: Interactions between non-native plant species and the floristic composition of common habitats, J. Ecol., 94, 1052–1060, https://doi.org/10.1111/j.1365-2745.2006.01172.x, 2006.

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., and Barr, C. J.: Vegetation Plots Handbook CS Technical Report No. 2/07, Centre for Ecology and Hydrology, Lancaster, 2008.

Maskell, L. C., Smart, S. M., Bullock, J. M., Thompson, K., and Stevens, C. J.: Nitrogen deposition causes widespread loss of species richness in British habitats, Glob. Change Biol., 16, 671–679, https://doi.org/10.1111/j.1365-2486.2009.02022.x, 2010.

Maskell, L. C., Crowe, A., Dunbar, M. J., Emmett, B., Henrys, P., Keith, A. M., Norton, L. R., Scholefield, P., Clark, D. B., and Simpson, I. C.: Exploring the ecological constraints to multiple ecosystem service delivery and biodiversity, J. Appl. Ecol., 50, 561–571, https://doi.org/10.1111/1365-2664.12085, 2013.

McCollin, D., Jackson, J., Bunce, R., Barr, C., and Stuart, R.: Hedgerows as habitat for woodland plants, J. Environ. Manage., 60, 77–90, https://doi.org/10.1006/jema.2000.0363, 2000.

Metzger, M., Bunce, R., Leemans, R., and Viner, D.: Projected environmental shifts under climate change: European trends and regional impacts, Environ. Conserv., 35, 64–75, 2008.

Metzger, M. J., Brus, D. J., Bunce, R. G. H., Carey, P. D., Gonçalves, J., Honrado, J. P., Jongman, R. H. G., Trabucco, A., and Zomer, R.: Environmental stratifications as the basis for national, European and global ecological monitoring, Ecol. Indic., 33, 26–35, https://doi.org/10.1016/j.ecolind.2012.11.009, 2013. Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R., and Simpson, I.: Final Report for LCM2007 – the new UK land cover map, NERC/Centre for Ecology & Hydrology, WallingfordCS Technical Report No. 9/07, 2011.

Norton, L., Inwood, H., Crowe, A., and Baker, A.: Trialling a method to quantify the "cultural services" of the English landscape using Countryside Survey data, Land Use Policy, 29, 449–455, https://doi.org/10.1016/j.landusepol.2011.09.002, 2012.

Norton, L., Greene, S., Scholefield, P., and Dunbar, M.: The importance of scale in the development of ecosystem service indicators?, Ecol. Indic., 61, 130–140, https://doi.org/10.1016/j.ecolind.2015.08.051, 2016.

Norton, L. R., Murphy, J., Reynolds, B., Marks, S., and Mackey, E. C.: Countryside Survey: Scotland results from 2007, Centre for Ecology & Hydrology, Wallingford, 2009.

Norton, L. R., Maskell, L. C., Smart, S. S., Dunbar, M. J., Emmett, B. A., Carey, P. D., Williams, P., Crowe, A., Chandler, K., Scott, W. A., and Wood, C. M.: Measuring stock and change in the GB countryside for policy: key findings and developments from the Countryside Survey 2007 field survey, J. Environ. Manage., 113, 117–127, https://doi.org/10.1016/j.jenvman.2012.07.030, 2012.

Ortega, M., Guerra, C., Honrado, J. P., Metzger, M. J., Bunce, R. G. H., and Jongman, R. H. G.: Surveillance of habitats and plant diversity indicators across a regional gradient in the Iberian Peninsula, Ecol. Indic., 33, 36–44, https://doi.org/10.1016/j.ecolind.2012.12.004, 2013.

Palmer, M. W.: Potential biases in site and species selection for ecological monitoring, Environ. Monit. Assess., 26, 277–282, https://doi.org/10.1007/BF00547504, 1993.

Petit, S., Griffiths, L., Smart, S. S., Smith, G. M., Stuart, R. C., and Wright, S. M.: Effects of area and isolation of woodland patches on herbaceous plant species richness across Great Britain, Landsc. Ecol., 19, 463–471, https://doi.org/10.1023/B:LAND.0000036141.30359.53, 2004.

Prosser, M. V. and Wallace, H. L.: Countryside Survey 1990 Quality Assurance Exercise. Report to Institute of Terrestrial Ecology, Merlewood Research Station, Canon Pyon, Herefordshire, 1992.

Prosser, M. V. and Wallace, H. L.: Countryside Survey 2000 Quality Assurance Exercise. Report to Institute of Terrestrial Ecology, Merlewood Research Station. Second draft, Canon Pyon, Herefordshire, 1999.

Prosser, M. V. and Wallace, H. L.: Quality Assurance Exercise, Report to Centre of Ecology & Hydrology, Lancaster, 2008.

Rhodes, C. J., Henrys, P., Siriwardena, G. M., Whittingham, M. J., and Norton, L. R.: The relative value of field survey and remote sensing for biodiversity assessment, Methods Ecol. Evol., 6, 772–781, https://doi.org/10.1111/2041-210X.12385, 2015.

Rodwell, J. S.: National vegetation classification: Users' handbook, Joint Nature Conservation Committee, Peterborough, 2006.

Rose, R., Monteith, D. T., Henrys, P., Smart, S., Wood, C., Morecroft, M., Andrews, C., Beaumont, D., Benham, S., and Bowmaker, V.: Evidence for increases in vegetation species richness across UK Environmental Change Network sites linked to changes in air pollution and weather patterns, Ecol. Indic., 68, 52–62, https://doi.org/10.1016/j.ecolind.2016.01.005, 2016.

Ross, L. C., Woodin, S. J., Hester, A., Thompson, D. B., and Birks, H. J. B.: How important is plot relocation accuracy when interpreting re-visitation studies of vegetation change?, Plant. Ecol. Divers, 3, 1–8, https://doi.org/10.1080/17550871003706233, 2010.

Rowe, E. C., Ford, A. E. S., Smart, S. M., Henrys, P. A., and Ashmore, M. R.: Using qualitative and quantitative methods to choose a habitat quality metric for air pollution policy evaluation, PLOS ONE, 11, e0161085, https://doi.org/10.1371/journal.pone.0161085, 2016.

Scott, W., Morecroft, M., Taylor, M., and Smart, S.: Countryside Survey-Environmental Change Network link, Centre for Ecology and Hydrology, Wallingford, UK, 2010.

Scott, W. A., Smart, S. M., and Clarke, R.: Quality Assurance Report: QA and bias in vegetation recording, Centre for Ecology and Hydrology, Lancaster, 2008.

Sheail, J. and Bunce, R. G. H.: The development and scientific principles of an environmental classification for strategic ecological survey in the United Kingdom, Environ. Conserv., 30, 147–159, https://doi.org/10.1017/S0376892903000134, 2003.

Silvertown, J., Poulton, P., Johnston, E., Edwards, G., Heard, M., and Biss, P. M.: The Park Grass Experiment 1856–2006: its contribution to ecology, J. Ecol., 94, 801–814, https://doi.org/10.1111/j.1365-2745.2006.01145.x, 2006.

Smart, S., Clarke, R., Van De Poll, H., Robertson, E., Shield, E., Bunce, R., and Maskell, L.: National-scale vegetation change across Britain; an analysis of sample-based surveillance data from the Countryside Surveys of 1990 and 1998, J. Environ. Manage., 67, 239–254, https://doi.org/10.1016/S0301-4797(02)00177-9, 2003.

Smart, S., Ashmore, M., Hornung, M., Scott, W., Fowler, D., Dragosits, U., Howard, D., Sutton, M., and Famulari, D.: Detecting the signal of atmospheric N deposition in recent national-scale vegetation change across Britain, Water Air Soil Poll., 4, 269–278, https://doi.org/10.1007/s11267-005-3037-5, 2005a.

Smart, S., Bunce, R., Marrs, R., LeDuc, M., Firbank, L., Maskell, L., Scott, W., Thompson, K., and Walker, K.: 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 changes in trait representation, Biol. Conserv., 124, 355–371, https://doi.org/10.1016/j.biocon.2004.12.013, 2005b.

Smart, S., Maskell, L., Dunbar, M., Emmett, B., Marks, S., Norton, L., Rose, P., and Simpson, I.: An Integrated Assessment of Countryside Survey data to investigate Ecosystem Services in Great Britain. Countryside Survey Technical Report No. 10/07, NERC Centre for Ecology and Hydrology, Lancaster, 2010a.

Smart, S., Henrys, P. A., Scott, W. A., Hall, J. R., Evans, C. D., Crowe, A., Rowe, E. C., Dragosits, U., Page, T., Whyatt, J. D., Sowerby, A., and Clark, J. M.: Impacts of pollution and climate change on ombrotrophic Sphagnum species in the UK: analysis of uncertainties in two empirical niche models, Clim. Res., 45, 163–177, https://doi.org/10.3354/cr00969 2010b.

Smart, S., Scott, W. A., Whitaker, J., Hill, M. O., Roy, D. B., Critchley, C. N., Marina, L., Evans, C., Emmett, B. A., Rowe, E. C., Crowe, A., and Marrs, R. H.: Empirical realised niche models for British higher and lower plants – development and preliminary testing, J. Veg. Sci., 21, 643–656, https://doi.org/10.1111/j.1654-1103.2010.01173.x, 2010c.

Smart, S. M. and DART Computing: Modular Analysis of Vegetation Information System (MAVIS), Centre for Ecology and Hydrology, http://www.ceh.ac.uk/services/modular-analysis-vegetation-information-system-mavis, last access: 16 November, 2017.

Smart, S. M., Firbank, L. G., Bunce, R. G. H., and Watkins, J. W.: Quantifying changes in abundance of food plants for butterfly larvae and farmland birds, J. Appl. Ecol., 37, 398–414, https://doi.org/10.1046/j.1365-2664.2000.00508.x, 2000.

Smart, S. M., Bunce, R. G. H., Firbank, L. G., and Coward, P.: Do field boundaries act as refugia for grassland plant species diversity in intensively managed agricultural landscapes in Britain?, Agr. Ecosyst. Environ., 91, 73–87, https://doi.org/10.1016/S0167-8809(01)00259-6, 2002.

Smart, S. M., Marrs, R. H., Le Duc, M. G., Thompson, K., Bunce, R. G., Firbank, L. G., and Rossall, M. J.: Spatial relationships between intensive land cover and residual plant species diversity in temperate farmed landscapes, J. Appl. Ecol., 43, 1128–1137, https://doi.org/10.1111/j.1365-2664.2006.01231.x, 2006a.

Smart, S. M., Thompson, K., Marrs, R. H., Le Duc, M. G., Maskell, L. C., and Firbank, L. G.: Biotic homogenization and changes in species diversity across human-modified ecosystems, P. Roy. Soc. B, 273, 2659–2665, https://doi.org/10.1098/rspb.2006.3630, 2006b.

Smart, S. M., Scott, A., and Clarke, R.: Quality Assurance Report – Analysis of plant species recording bias in Countryside Survey terrestrial vegetation plots – summary report, Centre for Ecology & Hydrology, Lancaster, 2008.

Smart, S. M., Allen, D., Murphy, J., Carey, P. D., Emmett, B. A., Reynolds, B., Simpson, I. C., Evans, R. A., Skates, J., Scott, W. A., Maskell, L. C., Norton, L. R., Rossall, M. J., and Wood, C.: Countryside Survey: Wales Results from 2007, NERC/Centre for Ecology & Hydrology, Wallingford, UK, 94, 2009.

Smart, S. M., Henrys, P. A., Purse, B. V., Murphy, J. M., Bailey, M. J., and Marrs, R. H.: Clarity or confusion?: problems in attributing large-scale ecological changes to anthropogenic drivers, Ecol. Indic., 20, 51–56, https://doi.org/10.1016/j.ecolind.2012.01.022, 2012.

Smart, S. M., Glanville, H. C., Blanes, M. d. C., Mercado, L. M., Emmett, B. A., Jones, D. L., Cosby, B. J., Marrs, R. H., Butler, A., Marshall, M. R., Reinsch, S., Herrero-Jáuregui, C., and Hodgson, J. G.: Leaf dry matter content is better at predicting above-ground net primary production than specific leaf area, Funct. Ecol., 31, 1336–1344, https://doi.org/10.1111/1365-2435.12832, 2017.

Ståhl, G., Allard, A., Esseen, P.-A., Glimskär, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Torell, Å. G., and Högström, M.: National Inventory of Landscapes in Sweden (NILS)—scope, design, and experiences from establishing a multiscale biodiversity monitoring system, Environ. Monit. Assess., 173, 579–595, 2011.

Stevens, C., Maskell, L., Smart, S., Caporn, S., Dise, N., and Gowing, D.: Identifying indicators of atmospheric nitrogen deposition impacts in acid grasslands, Biol. Conserv., 142, 2069–2075, https://doi.org/10.1016/j.biocon.2009.04.002, 2009.

Stevens, C. J., Payne, R. J., Kimberley, A., and Smart, S. M.: How will the semi-natural vegetation of the UK have changed by 2030 given likely changes in nitrogen deposition?, Environ. Pollut., 208, 879–889, https://doi.org/10.1016/j.envpol.2015.09.013, 2016.

van den Berg, L. J., Jones, L., Sheppard, L. J., Smart, S. M., Bobbink, R., Dise, N. B., and Ashmore, M. R.: Evidence for differential effects of reduced and oxidised nitrogen deposition on vegetation independent of nitrogen load, Environ. Pollut., 208, 890–897, https://doi.org/10.1016/j.envpol.2015.09.017, 2016.

Villoslada, M., Bunce, R. G. H., Sepp, K., Jongman, R. H. G., Metzger, M. J., Kull, T., Raet, J., Kuusemets, V., Kull, A., and Leito, A.: A framework for habitat monitoring and climate change modelling: construction and validation of the Environmental Stratification of Estonia, Reg. Environ. Change, 17, 1–15, https://doi.org/10.1007/s10113-016-1002-7, 2016.

Watson, R., Albon, S., Aspinall, R., Austen, M., Bardgett, B., Bateman, I., Berry, P., Bird, W., Bradbury, R., and Brown, C.: UK National Ecosystem Assessment: Technical Report, United Nations Environment Programme World Conservation Monitoring Centre, Cambridge, 2011.

Wood, C. M. and Bunce, R. G. H.: Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring, Earth Syst. Sci. Data, 8, 89–103, https://doi.org/10.5194/essd-8-89-2016, 2016.

Wood, C. M., Howard, D. C., Henrys, P. A., Bunce, R. G. H., and Smart, S. M.: Countryside Survey: measuring habitat change over 30 years: 1978 data rescue – final report, Centre for Ecology & Hydrology, Lancaster, 2012.

Wood, C. M., Smart, S. M., and Bunce, R. G. H.: Woodland Survey of Great Britain 1971–2001, Earth Syst. Sci. Data, 7, 203–214, https://doi.org/10.5194/essd-7-203-2015, 2015.

Earth Syst. Sci. Data, 10, 899–918, 2018
https://doi.org/10.5194/essd-10-899-2018

Open Access Earth System Science Data

Land cover and vegetation data from an ecological survey of "key habitat" landscapes in England, 1992–1993

Claire M. Wood[1], **Robert G. H. Bunce**[2], **Lisa R. Norton**[1], **Simon M. Smart**[1], **and Colin J. Barr**[a]

[1]Centre for Ecology and Hydrology, Lancaster Environment Centre, Bailrigg, Lancaster, LA1 4AP, UK
[2]Estonian University of Life Sciences, Kreuzwaldi 5, 51014 Tartu, Estonia
[a]formerly at: Institute of Terrestrial Ecology, Merlewood, Windermere Road, Grange-over-Sands, Cumbria, LA11 6JT, UK

Correspondence: Claire M. Wood (clamw@ceh.ac.uk)

Received: 21 November 2017 – Discussion started: 8 December 2017
Revised: 20 April 2018 – Accepted: 23 April 2018 – Published: 18 May 2018

Abstract. Since 1978, a series of national surveys (Countryside Survey, CS) have been carried out by the Centre for Ecology and Hydrology (CEH) (formerly the Institute of Terrestrial Ecology, ITE) to gather data on the natural environment in Great Britain (GB). As the sampling framework for these surveys is not optimised to yield data on rarer or more localised habitats, a survey was commissioned by the then Department of the Environment (DOE, now the Department for Environment, Food and Rural Affairs, DEFRA) in the 1990s to carry out additional survey work in English landscapes which contained semi-natural habitats that were perceived to be under threat, or which represented areas of concern to the ministry. The landscapes were lowland heath, chalk and limestone (calcareous) grasslands, coasts and uplands. The information recorded allowed an assessment of the extent and quality of a range of habitats defined during the project, which can now be translated into standard UK broad and priority habitat classes. The survey, known as the "Key Habitat Survey", followed a design which was a series of gridded, stratified, randomly selected 1 km squares taken as representative of each of the four landscape types in England, determined from statistical land classification and geological data ("spatial masks"). The definitions of the landscapes are given in the descriptions of the spatial masks, along with definitions of the surveyed habitats. A total of 213 of the 1 km^2 square sample sites were surveyed in the summers of 1992 and 1993, with information being collected on vegetation species, land cover, landscape features and land use, applying standardised repeatable methods. The database contributes additional information and value to the long-term monitoring data gathered by the Countryside Survey and provides a valuable baseline against which future ecological changes may be compared, offering the potential for a repeat survey. The data were analysed and described in a series of contract reports and are summarised in the present paper, showing for example that valuable habitats were restricted in all landscapes, with the majority located within protected areas of countryside according to different UK designations. The dataset provides major potential for analyses, beyond those already published, for example in relation to climate change, agri-environment policies and land management. Precise locations of the plots are restricted, largely for reasons of landowner confidentiality. However, the representative nature of the dataset makes it highly valuable for evaluating the status of ecological elements within the associated landscapes surveyed. Both land cover data and vegetation plot data were collected during the surveys in 1992 and 1993 and are available via the following DOI: https://doi.org/10.5285/7aefe6aa-0760-4b6d-9473-fad8b960abd4. The spatial masks are also available from https://doi.org/10.5285/dc583be3-3649-4df6-b67e-b0f40b4ec895.

1 Introduction

In Great Britain (GB), monitoring of ecological and land cover change has been carried out since 1978 via a programme named Countryside Survey (CS) (www.countrysidesurvey.org.uk, last access: 9 May 2018). The survey has been carried out at approximately decadal intervals by the Centre for Ecology and Hydrology (CEH) (and its predecessor, the Institute of Terrestrial Ecology, ITE) using quantitative and repeatable methods (Carey et al., 2008). The field survey uses 1 km squares as conveniently sized sampling units. These 1 km survey squares are a dispersed, stratified random sample, stratified using the "ITE Land Classification" (Bunce et al., 1996a), which is a statistical environmental classification of all 1 km squares in GB. Data from CS provides statistically robust national estimates of the quality and extent of a wide range of UK biodiversity action plan (BAP) broad habitats (Jackson, 2000) and also provides some estimates of the rarer priority habitats (Maddock, 2008). However, as the sampling framework for these surveys was designed to capture national estimates for ecological elements in the wider countryside, it is not optimised to yield data on features within rarer or more localised habitats.

In England, the former Department of the Environment (DOE) commissioned ITE (now part of CEH) to undertake a research project (Hornung et al., 1997) to quantify and evaluate the quality of the rarer semi-natural habitats of England not specifically covered by the more general monitoring CS provides, as a consequence of the widespread concern expressed over previous decades regarding the loss of semi-natural habitats, many of high nature conservation value. There has been considerable debate, particularly across Europe, about the relative importance of various drivers causing these losses, including changes in land use or farming practices, climate change, atmospheric pollution, or industrial and urban development.

Named as the "Key Habitat Survey" by ITE, the survey recorded vegetation species, land cover, landscape features and land use information from 1 km sample square sites occurring within the landscape types included as targets for conservation action in the original Countryside Stewardship Scheme (CSS) (Countryside Stewardship, 2017), an English grant scheme intended to reward farmers for farming land for nature conservation. The survey used established methods based on the standardised CS methods, as described below. In a variation to the standard CS methods, information was largely recorded from points falling on grid intersections within each 1 km square site, whereas, in CS, landscape area, point and line features are mapped across whole of 1 km square site, with vegetation plots being recorded at randomly placed points (i.e. not gridded) (Wood et al., 2017).

Standard habitat classes in Britain have evolved since the time of the Key Habitat Survey and can now be defined as broad (Jackson, 2000) and priority (Maddock, 2008) habitats. At the time of the project, a range of different customised land cover and habitat groupings were used to report the results of the survey. However, the data were recorded in such a way as to make it possible to translate information into the standard broad and priority habitat groupings, or European Annex I Classes as required (Romão, 2013). The "key habitats" term quoted in the title of the survey was derived from the term "key habitats" as included in the biodiversity action plans (UK Biodiversity Steering Group, 1995), which were later to evolve into the broad and priority habitat framework.

The surveyed landscape types were lowland heath landscapes, chalk and limestone (calcareous) grassland landscapes, coastal landscapes and upland landscapes. The main aims of the project were to determine the extent of a range of land cover types within each landscape type, to assess their ecological status and to establish a baseline for long-term monitoring of ecological change. All of the surveyed landscape types, together with their constituent broad and priority habitats, were seen as areas which had suffered serious losses and habitat degradation in the past and appeared to be still under threat. They were also perceived as having major significance for wildlife, landscape, archaeology and amenity criteria.

Information regarding specific habitats has become increasingly available through thematic and local surveys and inventories, such as Natural England surveys (Wilson et al., 2013; exegesis SDM Ltd. and Doody, 2009; Doody and Rooney, 2015; Jerram et al., 1998) and collation of information on lowland heath and calcareous grasslands (Marrs et al., 1986; Rose et al., 2000; Gibson and Brown, 1991; Moore, 1962). However, an important point is that the data from the Key Habitat Survey cover a range of the less common land cover and habitat types and offer an additional element to the long-term national monitoring programme of the Countryside Survey, both by providing additional data to augment the wealth of long-term ecological data already collected by the programme and by offering an additional targeted sampling framework, which could be incorporated into the Countryside Survey field survey should resources become available.

The data have hitherto remained unpublished, aside from the information in contract reports written following the field survey (Barr, 1996a, b, c, d). It is therefore timely that these data are now being made available for wider use.

2 The survey in context

There are a number of long-term national monitoring projects for widespread and more common habitats, particularly across Europe, for example in Switzerland (Hintermann et al., 2002), Norway (Dramstad et al., 2002) and Sweden (Ståhl et al., 2011), as well as globally (United States Forest Service, 2015; Wiser et al., 2001; Gillis et al., 2005). Local studies of specific habitats or specific species are also frequent in many countries, for example in Europe: peatlands in Slovakia (Špulerová, 2009), dunes in Belgium (Provoost et al., 2004), hay meadows in France (Broyer and Curtet, 2005), coastal monitoring in Ireland (Ryle et al., 2007) and other examples, which can be viewed in the EuMon database (EuMon, 2017). Beyond Europe, many other vegetation studies have also been undertaken, for example in Belize (Bridgewater et al., 2002) and Borneo (Aiba and Kitayama, 1999). In Britain, there are a range of examples of studies carried out in the last 50 years regarding the ecologically valuable landscapes covered by the Key Habitat Survey (Dargie, 1993, 1995; Radley and Dargie, 1994; Sneddon et al., 1994; Stevens et al., 2007). However, these studies specifically target individual habitat types, usually at a local level.

The Key Habitat Survey targeted a range of different, less common land cover and habitat types, contributing an additional element to the national ecological monitoring programme Countryside Survey, which provides a wide range of nationally significant ecological datasets, globally unique in their geographical coverage and time span. Other examples of structured, standardised, repeatable ecological data, targeted at a wide range of rare and localised habitats at a national level, are not known to the authors. The survey employs repeatable methods and is also designed in such a way as to add value to the Countryside Survey by offering additional targeted information regarding rarer and more localised habitats, which CS does not provide. The data regarding land cover, landscape features and vegetation collected during the survey offer detailed information with which to assess the quality and extent of the rarer broad and priority habitat types.

2.1 Landscape types

The Key Habitat Survey focused on the following landscapes: lowland heath landscapes, chalk and limestone (calcareous) grassland landscapes, coastal landscapes and upland landscapes. The choice of landscapes selected for survey was determined by their inclusion in the original Countryside Stewardship Scheme launched in 1991 in England. CSS was a grant scheme that offered payments to farmers and other land managers in order to make conservation part of normal farming and land management practice. The stated objectives of the scheme were to sustain the beauty and diversity of the landscape, improve and extend wildlife habitats, conserve archaeological sites and historic features, improve opportunities for countryside enjoyment, restore neglected land or features, and create new wildlife habitats and landscape features (Ovenden et al., 1998).

The lowland heath, calcareous and coastal landscapes are characterised to a greater or lesser extent by a mosaic of land cover types, including a variety of habitats. Thus, for example, lowland heath and calcareous grassland are the core broad and priority habitats occurring in the respective landscapes, but the landscapes also include many non-heath and non-calcareous grassland broad habitats (Jackson, 2000) (for example fen, marsh and swamp, neutral grassland and broadleaved woodland). Similarly, the upland and coastal landscapes include a range of habitats which are characteristically upland and coastal, in addition to other associated habitats.

The descriptions below highlight the importance of each landscape in containing broad, and particularly priority, habitats of high conservation value in a national, and in some cases international context, in addition to being valued scenically and recreationally.

2.1.1 Lowland heath landscapes

European heaths are widely recognised to be of high conservation value, as shown by their inclusion in Annex I of the EU Habitats Directive. The list includes Annex I habitats "4010: Northern Atlantic wet heaths with *Erica tetralix*", "4020: Temperate Atlantic wet heaths with *Erica ciliaris* and *Erica tetralix*", "4030: European dry heaths" and "4040: Dry Atlantic coastal heaths with *Erica vagans*" (Romão, 2013). Lowland heath occurs across continental Europe, but the British heaths are especially important in conservation terms, in part because they form such a large proportion of the European resource. For example, Farrell (1989) estimated that Britain contains 18 % of the total European resource, including wet heath and maritime heath vegetation types which are relatively rare. In the UK, lowland heath was designated as a priority habitat under the national biodiversity action plan, reflecting its rare and threatened status (Maddock, 2008), as well as its importance for a number of characteristic species of vascular plants, bryophytes and lichens, supporting characteristic birds, reptiles, amphibians and invertebrates (Department of the Environment, 1995).

The distribution of the lowland heath landscapes is largely controlled by particular combinations of geology and soils with lowland heath occurring on acidic, often podzolic soils that are low in nutrients, mainly as a result of soil deterioration in prehistoric times. However, important bog and wet heath habitats in the lowland heath landscape are associated with wetter acid soils.

Lowland heaths have become the focus of increasing conservation concern as a result of high rates of loss and degradation. For example in Sweden and Denmark, the area of this habitat declined by 60–70 % in the century prior to the 1960s, with the corresponding decline for the Netherlands being 95 % (Farrell, 1989). The survival of the distinctive lowland heath vegetation and habitats, dominated by heather (*Calluna vulgaris*) and gorse (*Ulex europaeus*), is dependent on traditional use, including livestock grazing, cutting of the shrub for use as fuel and animal fodder, or controlled burning (Dolman and Land, 1995). Much of the decline and fragmentation of heaths is attributable to changing patterns of land use, including agricultural intensification, afforestation, mineral extraction and urban development (Webb, 1986). As a result of these factors, many heaths have reverted to scrub or woodland through a process of natural succession, or have been converted into intensive grassland. In the UK, the extent of lowland heaths is now approximately one sixth of that present in 1800 (Department of the Environment, 1995). The decline of the Dorset heaths has been especially well studied (for example, Moore, 1962; Pywell et al., 1997; Rose et al., 2000); the area has dropped from around 40 000 ha in 1760 to 5800 ha by 1978 (Webb and Haskins, 1980). Today most areas of lowland heath are used for low intensity grazing, military training and recreation, with some areas in the latter two categories areas being unmanaged.

In England, the largest remnants are concentrated in the New Forest, Breckland, the Suffolk Sandlings, East Hampshire, Surrey, Dorset and the Lizard.

2.1.2 Calcareous landscapes

Calcareous grasslands are associated with shallow, calcareous soils overlying limestone and chalk bedrock. The type of grassland varies with the type of underlying calcium-rich bedrock, with the principle division being between the chalk grasslands on soft substrates in the south and east of England and the limestone grasslands occurring on harder Carboniferous strata in the north and west of Britain.

Calcareous grasslands are botanically rich, being amongst the most species-rich and species-diverse plant communities in Britain and northern Europe. In Annex I of the EU Habitats Directive, the following are included: "6210/6211, Semi-natural dry grasslands and scrubland facies on calcareous substrates (Festuco-Brometalia) (including important orchid sites)". Within Britain, the large number of plant species occurring in calcareous grassland constitutes a substantial percentage of the total native flora (estimated at 10–20 %) and many of the plant species are scarce native species; a total of 77 protected or listed species occur in calcareous grassland, of which 50 are restricted to calcareous grassland only (Keymer and Leach, 1990). In addition, calcareous grasslands (especially on the warm South Downs) provide habitats for many invertebrates including ants and butterflies which are confined to this region and are scarce or localised in Britain. In contrast to lowland heaths, England only contains a small part of the European stock of calcareous grassland; such grasslands occur over much of central and northern Europe. However, their rarity in Britain makes them a nationally important resource and they are listed as priority habitats "upland calcareous grassland" and "lowland calcareous grassland" (Maddock, 2008).

The extent of calcareous grassland is thought to have reached a maximum 300 years ago. Since then, large areas have been lost, with substantial losses occurring within the last 70 years (Poschlod and WallisDeVries, 2002; Fuller, 1987). The introduction of seeding agricultural grassland after 1700 led to a decline in the quality of some chalk grassland, and as farming became mechanised in the early 19th century, many grasslands were ploughed up. During the 20th century many calcareous grasslands have been lost to arable or improved pasture, mineral extraction, afforestation and building development. Keymer and Leach (1990) suggested that between 1968 and 1980 the loss of grassland was about 60 % due to ploughing or agricultural improvement, about 30 % to scrub encroachment and 1 % due to development. As most calcareous grassland remains in agricultural ownership, the impact of changes in agricultural management is significant and grazing is the dominant influence in the maintenance of calcareous grassland. In England, the largest areas are in the south, such as Salisbury Plain, and the North and South Downs. They also occur in Yorkshire, Derbyshire, Morecambe Bay and County Durham.

2.1.3 Coastal landscapes

Coastal habitats and land cover types tend to be dynamic compared to those in the other surveyed landscapes. Geology is a major factor determining the type of coastal landscape and the constituent habitats, with the major division being between soft and hard rock coasts, with the former associated with salt marshes and low earth cliffs and the latter with rocky foreshores and cliffs. Within these major divisions there is a mosaic of habitat types. Early successional plant communities are particularly important in the coastal zone, in comparison to the other landscapes. Many of the habitats in the coastal landscape are of restricted occurrence and contain rare species. Stewart et al. (1994) estimate that at least 20 % of the nationally scarce plants (Joint Nature Conservation Committee, 2018) in Britain are coastal. Coastal habitats listed as priority habitats in the UK biodiversity action plan (Maddock, 2008) include coastal and floodplain grazing marsh, coastal salt marsh, coastal sand dunes, coastal vegetated shingle, maritime cliff and slopes, and intertidal mudflats. The UK has special responsibility (as it holds a large proportion of the European resource) for several coastal habitats listed in the EU Habitats Directive, including "1230: Vegetated sea cliffs of the Atlantic and Baltic coasts", "1160: Large shallow inlets and bays" and "1130: Estuaries".

Coastal landscapes have often been heavily influenced by man, although some of the core maritime habitats are formed naturally. The coastal belt is particularly well used for a wide variety of recreational activities. The detailed mix of species and the mosaic of habitats (including cliffs, estuaries, mudflats and beaches) are inevitably influenced by the management and use of the landscapes.

2.1.4 Upland landscapes

In the uplands, the interaction between the underlying soils, geology and climate determine the collection of habitats which make up the landscape. This landscape occurs largely in the north of the country, extending from Northumberland to the Pennines, Yorkshire Dales, Derbyshire and Lake District, but with important outliers in the south-west, notably Dartmoor and Exmoor.

The combination of montane and oceanic climatic conditions gives rise to plant communities which are of restricted distribution in Europe. Although the habitats are relatively species poor, they are often present as large continuous units extending over extensive expanses of land, which are rare elsewhere in Britain. They therefore support species of birds that might not persist in smaller, more fragmented habitats, such as hen harriers (*Circus cyaneus*), merlin (*Falco columbarius*) and raven (*Corvus corax*), as well as breeding waders (Thompson et al., 1995; Usher and Thompson, 1993). Upland priority habitats include upland heaths, upland flushes and blanket bog. Upland habitats listed in the EU Annex I directive include "7130: Blanket bogs", "4060: Alpine and Boreal heaths" and "4030: European dry heaths".

Much of the upland landscape has been dominated by upland heaths and bogs since the Iron Age (Tallis, 1991). It would also have been forested at some point since the last glacial period. Whilst management, grazing and burning are important in maintaining the mix of habitats in the uplands, it is not likely that reversion to scrub or woodland would occur in all the formerly wooded areas, due to peat formation and the current climate.

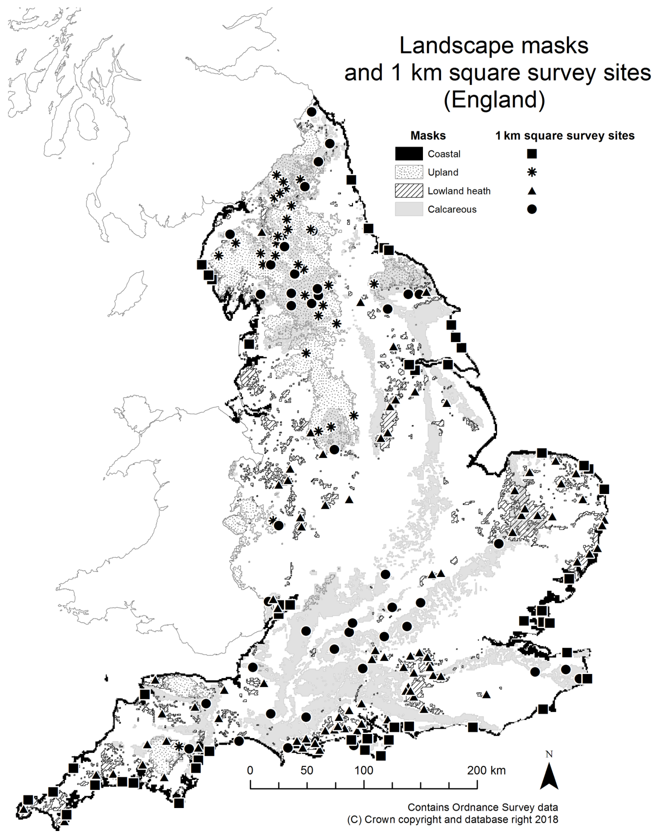


Figure 1. Distribution of spatial landscape masks and 1 km square survey sites.

3 Survey design: 1 km square site selection and stratification

The overall design of the Key Habitat Survey, in principle, follows the standardised procedures described by Bunce and Shaw (1973). The methods are utilised in the national Countryside Survey 1978–2007 (Carey et al., 2008) and also the recent Welsh Glastir Monitoring and Evaluation Programme 2013–2016 (Emmett and GMEP team, 2017). The methods have also been successfully deployed in a range of British regional surveys (Wood and Bunce, 2016; Bunce and Smith, 1978; Wood et al., 2015). A comparison of the sampling approaches used in both the Countryside Survey and the Key Habitat Survey is given in Table 1.

In the same way to CS, the Key Habitat Survey uses a sampling approach, with random samples of 1 km squares being selected for survey from a statistical environmental classification to enable robust estimates of areas to be produced. This stratified, random strategy ensures adequate representation of the range of ecological variation within the landscapes. Whereas CS uses the ITE Land Classification to form a sampling framework for the GB Countryside Survey (Bunce et al., 1996a), the Key Habitat Survey uses a more targeted set of "spatial masks" to stratify the samples for each landscape type (incorporating the ITE Land Classification to some extent). The ITE Land Classification was initially developed in the late 1970s and uses a range of environmental variables such as altitude, climate, geology, human geography and location. Using multivariate analysis, GB was split into a set of 32 land classes (or strata), from which the 1 km survey squares could be randomly selected.

In terms of the Key Habitat Survey, only fragmentary information existed at the start of the project from which to define and map the national distribution of the landscapes. Procedures were therefore developed to create a mask for each landscape which defined those 1 km squares in England, which contained, or had the potential for containing, the characteristic habitats of that particular landscape, thus providing the environmental classification required for the stratification framework (Fig. 1 and Table 2). Nature sites and areas of countryside can be "designated", which means they have special status as protected areas because of their natural and cultural importance (Government Digital Service, 2018). Additional information regarding UK designation (designated or non-designated) (Natural England, 2017a) was also utilised to facilitate the choice of 1 km survey squares. In this context designated refers to the following: site of special scientific interest (SSSI), national nature reserve (NNR), national park (NP), area of outstanding natural beauty (AONB), heritage coast (HC), green belt, and environmentally sensitive areas (ESA). The 1 km sample squares were drawn at random from within the landscape masks and randomly sampled (Fig. 1) with land cover, vegetation in quadrats and landscape ele-

Table 1. Table showing a comparison between the Countryside Survey and the Key Habitat Survey.

	Countryside Survey	Key Habitat Survey
Coverage	Great Britain	England
Periodicity	5 (between 1978 and 2007)	1 (to date)
Information collected	Vegetation plots, landscape feature mapping, soils, freshwater	Vegetation plots, landscape feature mapping
Sampling design	Stratified random sampling	Stratified random sampling
Sampling stratification	ITE Land Classification	Geographical landscape masks (see Sect. 3)
In-square sampling methodology (feature mapping)	Mapping of all point, line and areas features in square	Grid-based mapping; features mapped at grid intersections
Vegetation plot locations	Predetermined dispersed random sampling plus targeted plots	Predetermined dispersed random sampling (plus some targeted plots), largely based on a gridded design
Sampling intensity	$591 \times 1\,km^2$	$213 \times 1\,km^2$
Sampling unit	1 km squares	1 km squares
Optimised for:	Common and widespread habitat types	Rare and localised habitat types

ments being recorded in field surveys. Historic features were also recorded but are beyond the scope of this paper. The location of the vegetation quadrats was permanently marked to facilitate resurvey. In total, 213 squares were surveyed across England.

3.1 Defining the lowland heath mask

The lowland heath landscape mask contains existing and potential areas of what could now be classed as the priority habitat, "lowland heath". The mask was constructed by combining data on soils and altitude. Soil types characteristic of lowland heath vegetation and landscapes were used to define a population of 1 km squares having potential for heath. A 1 km dataset of the Soil Survey and Land Research Centre (Cranfield University, 2017) provided data in digital form on dominant and sub-dominant soils within 1 km grid squares. Soil types most likely to support heath vegetation were identified, along with the soil types appearing in areas of known heaths. Peat soils were also included as these have a potential for heaths, especially in the vicinity of existing heathland. A full list of soil types used is given in the supporting documentation accompanying the dataset.

Soils data alone cannot be used to differentiate between upland and lowland heaths, and neither can lowland heath simply be defined in terms of altitude. As climate varies in different parts of England, that which might be considered upland vegetation in some places may occur at relatively low altitudes in harsher environments. Thus, whereas the lowland–upland vegetation interface may be considered to occur somewhere in the region of 200–300 m in the south of England, in the north characteristically upland vegetation may occur in areas around sea level. In order to overcome these regional differences, we made use of the ITE Land Classification 1990 (Bunce et al., 1990). This consists of a statistical environmental classification covering the whole of Great Britain, created by the multivariate analysis of environmental factors, for example altitude and climate, from each 1 km square in the country (Bunce et al., 1996b). This classification used a range of environmental and physical parameters to assign all the 1 km squares in Great Britain into one of 32 land classes; land classes 17–24 and 27–28 which are characteristically upland in nature were used to exclude areas of England unlikely to contain lowland heath landscape areas. Coastal heathlands are poorly covered by this mask because they tend to be small and difficult to associate with soil types marked on the 1 : 250 000 soil map. Attempts were made to identify soils in areas of known coastal heathlands so that they could be incorporated into the lowland heath mask; however, the soils identified were not specific to coastal heathland areas and no procedure could be devised to limit the soil types to those areas. However, coastal heathlands are part of the coastal mask. The lowland heath mask covers $8538\,km^2$ in lowland England.

3.2 Defining the calcareous grassland mask

The calcareous grassland landscape mask covers $26\,555\,km^2$ in England, containing existing and potential areas of what can now be classed as the broad habitat, "calcareous grassland". Areas of potential calcareous grassland were identified by using a combination of data on solid (bedrock) geology and quaternary deposits. Simplified digitised versions of the 1 : 625 000 British Geological Survey (BGS) solid geology and quaternary maps (drift geology) of Britain were employed (British Geological Survey, 2017). Using these data, a 1 km resolution map was defined by identifying 1 km squares dominated by marine limestones, oolitic and friable limestones, and metamorphic limestones, excluding squares where the rocks are overlain with non-calcareous soils. Any adjacent 1 km squares containing steep slopes were added to improve the coverage of calcareous areas found on escarpments. Squares with more than 75 % urban land were excluded.

Table 2. Summary of the spatial landscape mask definitions. SSLRC is the Soil Survey and Land Research Centre

Definition of the landscape	Data used in defining the mask	Characteristic broad habitats	Characteristic priority habitats
Lowland heath			
Containing habitats with vegetation consisting of calcifuge species usually with dwarf shrubs and often containing species of southern distributions	Distribution of soil types characteristic of lowland heath (SSLRC 1 km dominant soil type; Cranfield University, 2017) ITE land classes 1–16, 25 and 26 (lowland) excluding upland classes 17–24 and 27–28 (Bunce et al., 1990)	Dwarf shrub heath	Lowland heath
Calcareous grassland			
Containing habitats with vegetation having a major component of calcicole species, often containing rare plants	Distribution of limestone and chalk bedrock excluding areas overlain with drift deposits (British Geological Survey, 2017) Adjacent 1 km squares containing steep slopes to ensure inclusion of limestone escarpments	Calcareous grassland	Upland calcareous grassland Lowland calcareous grassland
Coastal landscape			
Containing coastal habitats having vegetation usually with a major component of maritime species, with some exceptions such as nitrophilous patches	All land within 500 m of the coastline as defined on the Land Cover Map 1990 (Fuller et al., 1993), plus any contiguous areas of coastal vegetation (sand dunes, shingle and salt marsh) extending seaward of this coastal zone	Supra-littoral rock Supra-littoral sediment Littoral sediment	Maritime cliffs and slopes Sand dune Strandline/coastal vegetated shingle Coastal salt marsh
Upland landscape			
Contains upland habitats defined as having vegetation usually consisting calcifuge species and bog plants, often with dwarf shrubs and with local patches of arctic-alpine plants	ITE land classes 17–24 plus 27–28, the English land classes considered to be primarily upland in character (Bunce et al., 1990)	Acid grassland Bracken Dwarf shrub heath Fen, marsh, swamp Bog	Purple moor grass rush pasture Blanket bog Upland heath

3.3 Defining the coastal mask

The coastal landscape mask was defined as that area of land extending 500 m inland from the mean high water mark (HWM) plus all contiguous areas of salt marsh, dunes and coastal bare land. The 25 m resolution Land Cover Map 1990, a satellite-derived map of UK land cover types (Fuller et al., 1993), gave the location of the HWM and this was chosen for use. A coastal buffer was defined as a set of contiguous 1 km grid cells in England where coastal attributes (i.e. coastal buffer, salt marsh or coastal bare) were present. In total, 8870 km squares were covered in some part by the coastal zone. Of these, 787 urban squares (> 75 % built up) and 742 squares which were predominantly at sea were also excluded, leaving a total of 7341 km squares in England. The coastal mask was further sub-divided into estuarine, soft and hard coasts. As the coastal areas are narrow zones around the coast, squares often contain a proportion of the sea.

Table 3. Data collected regarding land cover and area features.

Attribute	Description
Theme	Broad land use category, e.g. agricultural crops
Primary attribute	Feature name, e.g. potatoes
Use	Use category, e.g. hay, timber production
Species	Species where relevant
Species cover	Cover of above species across polygon
Heights	Height of plants
Primary qualifiers	Additional information pertaining to primary attribute, e.g. number of horses, canopy descriptions, "windblow"
Age	Approximate age of tree species
Management	Description of land management, e.g. abandoned, mown

3.4 Defining the upland mask

Again, it was not adequate to simply define the upland landscape by altitude alone. To allow for the inherent variation in land above certain altitudes in different parts of England, the upland mask was derived from the ITE Land Classification 1990 (Bunce et al., 1990), as this stratification provides an overall integration between the critical environmental factors. As described above, the predominantly upland classes include 17–24 and 27–28 and thus were used as the basis of the mask. Squares which were predominantly urban (51) were excluded, providing a mask area of 15 616 km^2.

4 Data collected

The lowland heath landscapes were surveyed in the summer of 1992, with the remaining three landscape types surveyed in 1993. In a variation to the Countryside Survey methodology (Maskell et al., 2008a, b), information was collected based on a grid-based sampling framework within each 1 km square survey site, as shown in Fig. 2. Coastal and lowland heath landscapes used a 25-point grid, and calcareous and upland landscapes used a 16-point grid. Grid points were marked on base maps and located in the field using measurements and bearings from prominent features.

Rules were in place for relocating points falling on linear features or in urban land. The detailed rules for relocation are given in the field handbooks (Barr, 1992, 1993), although the general rule meant moving the point 10 m away from the original grid point where possible.

With maximum resource, the ideal survey methodology would follow exactly the methods of the Countryside Survey as described in Wood et al. (2017, 2018) in order to obtain the most comprehensive dataset for a full understanding of the landscapes in question. In terms of the land cover and boundary data, this would mean that the whole of each 1 km survey square site would be fully mapped with landscape point, line and area features. Whilst the grid-based approach has the potential to save time in the field, much information regarding structure and pattern is lost. A further assessment of alternative methods is described in Wood et al. (2018). In terms of the vegetation data, the approach taken has been proven as being highly effective for assessing the quality of vegetation at a national scale, as described in Wood et al. (2017).

4.1 Land cover data

4.1.1 Land cover data: areas

Land cover at each grid point in each square was described using a comprehensive list of land use and land cover codes, as used in Countryside Survey 1990 (Barr, 1990). Recorded attributes are summarised in Table 3. All mappable units included a primary description of the feature in question (for example maritime grassland, fen, scrub), along with dominant species (> 25 %) and percentage cover codes, as well as use or other descriptive codes where appropriate (for example cattle, hay). A full list of these codes can be found in the field survey handbooks (Barr, 1992, 1993), supplied as supporting information with the datasets. The codes reflected the mappable unit, or patch, in which the point fell. The minimum mappable unit (MMU) was 400 m^2. Each patch defined was determined by the constancy of the descriptive codes within. If one characteristic (e.g. cover of a dominant plant species) was different from that in an adjacent area, a different code was required and a new patch was distinguished. It is possible to allocate features to standard groupings, for example the broad habitat classification. Table S1 in the Supplement indicates the broad (and in some cases priority) habitat allocations for the mapped field codes.

4.1.2 Land cover data: boundaries

The nearest vertical boundary (measuring > 20 m in length) to each grid point in each square (within 100 m) was described using codes, as used in Countryside Survey 1990. Codes included a primary description of the feature (or combination of features) in question (for example "fence", "hedge" "earth/stone bank"), along with heights, an assessment of quality (for example "stock proof", "derelict"), and dominant species and percentage covers (in hedges or lines of trees). A full list of these codes can be found in the field survey handbooks (Barr, 1992, 1993), which is supplied with

Table 4. Data collected regarding boundaries.

Attribute	Description
Theme	Feature name, e.g. bank, inland water, woody linear feature
Primary attribute	Feature type, e.g. stone bank, canal
Height	Height of feature, where appropriate
Age	Approximate age of tree species
Evidence management	Evidence of recent management, e.g. cutting, flailing
Staked trees	Staked individual trees within the feature
Tree protectors	Tree protectors
Stock proof/gaps	Whether feature is stock proof; assessment of gaps in feature
Species	Tree/shrub species
Proportion	Proportion of species in feature

the data. Recorded attributes are summarised in Table 4. The point on the boundary which was nearest to the grid point was recorded as part of a length which could be coded constantly as part of a single unit of not less than 20 m (the minimum mappable length, MML). If the nearest point on the boundary was part of a longer length, then the coding reflected the variability of the longer length.

4.2 Vegetation data

Sampling of vegetation from within quadrats (i.e. plots) largely used the methodology followed by the Countryside Survey (Wood et al., 2017) with variations as detailed below. At each plot, the slope, aspect, shade, general soil type and descriptions were recorded. A summary of the number and locations of plots recorded is given in Tables 5 and 6.

In each plot, a complete list of all vascular plants and a selected range of readily identifiable bryophytes and macrolichens was made. The field training course held before the surveys covered identification of difficult species, regular visits were made to survey teams by managers, and difficult specimens could be collected and sent to experts for identification. Cover estimates were made to the nearest 5 % for all species reaching at least an estimated 5 % cover. Presence was recorded if cover was less than 5 %. Predetermined combinations of species may have been recorded as aggregates reflecting known difficulties in their separation in the field (refer to Barr, 1993).

4.2.1 X plots

The term "X plot" is used to denote plots located at predetermined, dispersed random sampling points. In this survey, 2 different sizes of X plot were used, 4 and 200 m^2, as described below.

X plots – 4 m^2

These small plots were only recorded in the lowland heath and calcareous landscape types. In lowland heath landscapes, a 4 m^2 X plot was located at each of 25 points on the grid (Fig. 2). In calcareous landscapes, five of these plots were located at points "A", "J", "G", "D" and "P" (see Fig. 2). Points were pre-marked on base maps and were laid out with the map point forming the south-east corner of the plot. Using canes and measuring tapes, a square with sides of 2 m in length was measured out and was oriented north–south.

X plots – 200 m^2

These large, 200 m^2 (14.14 × 14.14 m) plots were used in 1993 in the coastal and upland surveys. Five plots were placed at pre-selected randomised points on a grid within the squares. The rules for the placement of these plots were as follows: in coastal squares, X plots were recorded where possible at points "A", "L", "I", "T" and "W" on the 25-point grid (see Fig. 2). In upland squares (16-point grid), the X plots were recorded at "A", "J", "G", "D" and "P". Where land at the intersection in question was built-up, a lake, road, railway line, river or sea (below low water mark, LWM), and then another point was selected, with the nearest northern point being chosen first, rotating clockwise. X plots in arable fields or highly improved grassland were not recorded.

The methodology for 200 m^2 X plots was originally produced for woodlands as described by Bunce and Shaw (1973) and was also used and found appropriate for strategic ecological surveys (Bunce and Smith, 1978). The design of the plot not only aids a systematic search of the vegetation present but ensures a standard area of the plot is covered on every occasion. The plot is set up by using a centre post and four corner posts, with a set of four strings tagged with markers at specified distances. The tagged strings form the diagonals of the square. The diagonals are orientated carefully at right angles with the strings on the north–south, east–west axes. Within the each plot, the initial nest (2 × 2 m) is searched first. This procedure is then repeated for each nest of the quadrat, increasing the size each time and only recording additional species discovered in each larger nest. In the final nest (the whole 200 m^2 plot), the percentage cover (to the nearest 5 %) of each species is also estimated. Estimates of cover for litter,

Table 5. Summary of vegetation plot locations.

Landscape type	No. of 1 km squares	Map grid	*X* plots (200 m^2)	*X* plots (4 m^2)	*Y* plots (4 m^2)	S/W plots (10 × 1 m)	R/V plots (10 × 1 m)	Year surveyed
Lowland heath	89	25 points, A–Y	–	25 plots, on grid	–	–	–	1992
Calcareous	43	16 points, A–P	–	5 plots recorded at AJGDP	5 at locations selected by surveyor	–	5 plots adjacent to roadsides	1993
Coastal	49	25 points, A–Y	5 plots recorded at points ALITW	–	5 at locations selected by surveyor	–	–	1993
Upland	32	16 points, A–P	5 plots recorded at AJGDP	–	5 at locations selected by surveyor	5 plots adjacent to watercourses	–	1993

Table 6. Summary of vegetation plots recorded.

Landscape type	No. of 1 km squares	X plots (200 m^2)	X plots (4 m^2)	Y plots (4 m^2)	S/W plots (10 × 1 m)	R/V plots (10 × 1 m)
Lowland heath	89	–	553	–	–	–
Calcareous	43	–	122	215	–	81 (R) 120 (V)
Coastal	49	92	–	245	–	–
Upland	32	148	–	160	60 (S) 90 (W)	–
Total	213	240	675	620	150	201

wood, rock and bare ground are also included where present. Vegetation height, aspect and slope are also recorded. This approach is to ensure that the whole plot is observed consistently and systematically, avoiding unstructured search routines which are more likely to lead to species being overlooked, as described as far back as 1940 by Hope-Simpson (1940). The method has been widely tested and shown to be robust, not only in resource assessment, but also in measuring change.

4.2.2 Y plots 4 m^2

Five of these small targeted plots were placed in each square in semi-natural vegetation types that were not covered by the main (X) plots. These type of plots were used in 1993, in the coastal, upland and calcareous surveys. The five plots were placed randomly in five different land cover types where available, additional to those types already represented by the five randomly located (X) plots. If there were more than five land cover types available, priority was given first to those most typical of the landscape type, and second to the size of the area in question. If there were fewer than five land cover types, plots were placed proportionally to the number of land cover types available. These Y plots were important in sampling fragments of semi-natural habitat particularly in lowland landscapes, where patches may be small and embedded in a matrix of intensive farmland. Of all the plots recorded, they are most similar to the approach taken when positioning relevés (quadrats) during national vegetation classification (NVC) (Rodwell, 2006) because their location is not predetermined.

4.2.3 S/W plots – streamside plots

Up to five of these linear (10 × 1 m) plots were placed immediately adjacent to watercourses where present, in the upland landscapes only (in 1993). The term streamside plot denotes linear plots which lie alongside running water features (mainly rivers and streams, but also canals and ditches). Two streamside (S) plots were established, located as close as possible to the two large X plots in each square, which were furthest apart. Up to three additional Waterside (W) plots, representing other waterside types were included where appropriate.

4.2.4 R/V plots – roadside and verge plots

Up to five of these linear (10 × 1 m) plots were placed immediately adjacent to roads where present – in the calcareous landscapes only. The term "roadside plot" denotes those linear plots which lie alongside transport routes (mainly roads and tracks). The "R" and "V" prefixes refer to the different origins of the plots within the Countryside Survey: two roadside (R) plots were established, located as close as possible to the two X plots in each square, which were furthest apart. Up to three additional verge (V) plots were placed in verges alongside other transport routes where present in the square.

5 Data quality and repeatability

5.1 Spatial landscape masks

Work was carried out to validate the masks (mainly the calcareous and lowland heath) through comparisons with other datasets, although none of these provided definitive or directly comparable data for validation purposes. As the coastal and upland masks were more straightforward to define geographically, and the best available relevant data (at the time) were used in defining the masks, comparisons with other data were therefore not appropriate. The calcareous mask was compared against soils data (Mackney et al., 1983), as well as the former English Nature (EN) database on calcareous sites (Natural England, 2017b). The lowland heath was compared to the satellite-derived Land Cover Map 1990 (Fuller et al., 1993) and to English Nature lowland heath sites (Natural England, 2017b). Overall, the lack of resolution resulting from the use of the 1 km square geological data caused some discrepancies in comparison with these other datasets. However, at the time, this was the only geological dataset available for use in the project (higher-resolution geological data were in existence). In terms of the calcareous mask, the match with the English Nature data was good, covering 89 % of the EN chalk sites and 87 % of the EN limestone sites. The lowland heath mask covered only 55 % of the lowland heathland sites registered by English Nature. Most of the sites not covered by the lowland heath mask are scattered throughout England, but there is a particularly poor coverage in areas of Hampshire and Cornwall. In these areas, the missing sites occur on 1 km squares with dominant or subdominant soil types which are not specific to lowland heathland, and it was not possible to improve the coverage of the lowland heath mask without greatly increasing its size to cover large areas of England with little or no heathland potential. The map of lowland heathland areas derived using only soils and land class data therefore missed many small pockets of heathlands. However, with the exception of coastal heathlands, and areas in the New Forest and Cornwall where there are several mismatches between the Land Cover Map 1990 and English Nature's reference database and the lowland heathland map, most areas of existing heathlands were adequately covered.

The overall conclusion was that, although there were some mismatches between the masks and other datasets, the fit was judged to be acceptable for the purposes of the project in providing an adequate sampling framework. Whilst it is acknowledged that with the increased quality and availability of digital data the masks could be improved, a key aim of the sampling framework (heavily based on the ITE Land Classification) is that it provides an objective and static sampling framework, independent of specific environmental indicators being measured. As the underpinning data used in the classification is static over time, the classes themselves will not change and repeat surveys and repeat analyses are possible and easily comparable. The consistency in sampling protocols is crucial for robust, repeat analyses.

5.2 Field survey data

Several approaches were used to maintain quality in field recording and to minimise variation between surveyors. The field surveys were carried out by teams of experienced botanical surveyors and were preceded by intensive training courses, ensuring high standards and consistency of methodology, effort, identification and recording across the survey according to criteria laid out in the field handbooks (Barr, 1992, 1993). During the surveys, survey teams were initially supervised and later monitored by experienced project staff in order to control data quality. Data were recorded on waterproof paper sheets and were consequently transferred from the original field sheets to spreadsheets, using a "double-punch" method to minimise errors in data entry. They were checked using range and format checks and corrected to produce a final validated copy.

During the field survey, independent ecological consultants revisited a sample of the survey squares and repeated quadrats and land cover descriptions. The unpublished results show a 74.3 % accuracy rate in the recording of vegetation plots, comparable to the CS 1990 accuracy of 74–83 % (Prosser and Wallace, 1992). Information from these repeat visits was given to surveyors so that consistency of recording was maintained.

5.2.1 Plot relocations

During the surveys, plot locations were recorded on paper using a sketch map with measurements from distinguishing landscape features and by taking at least two photographs, preferably also including key landscape features in proximity to the plot. In addition to these, permanent metal plates or wooden stakes were placed in the ground to mark the plot locations. These steps were taken in order to facilitate any potential future visits to the plots.

The methods used to mark plots are identical to the methods used in Countryside Survey which have been widely tested and shown to be robust. The CS plots are estimated to have a precise relocation accuracy of 85–86 % (Prosser and Wallace, 2008), and, in the event of a resurvey of these key habitat plots, it would be expected that the plot relocation accuracy would be similar.

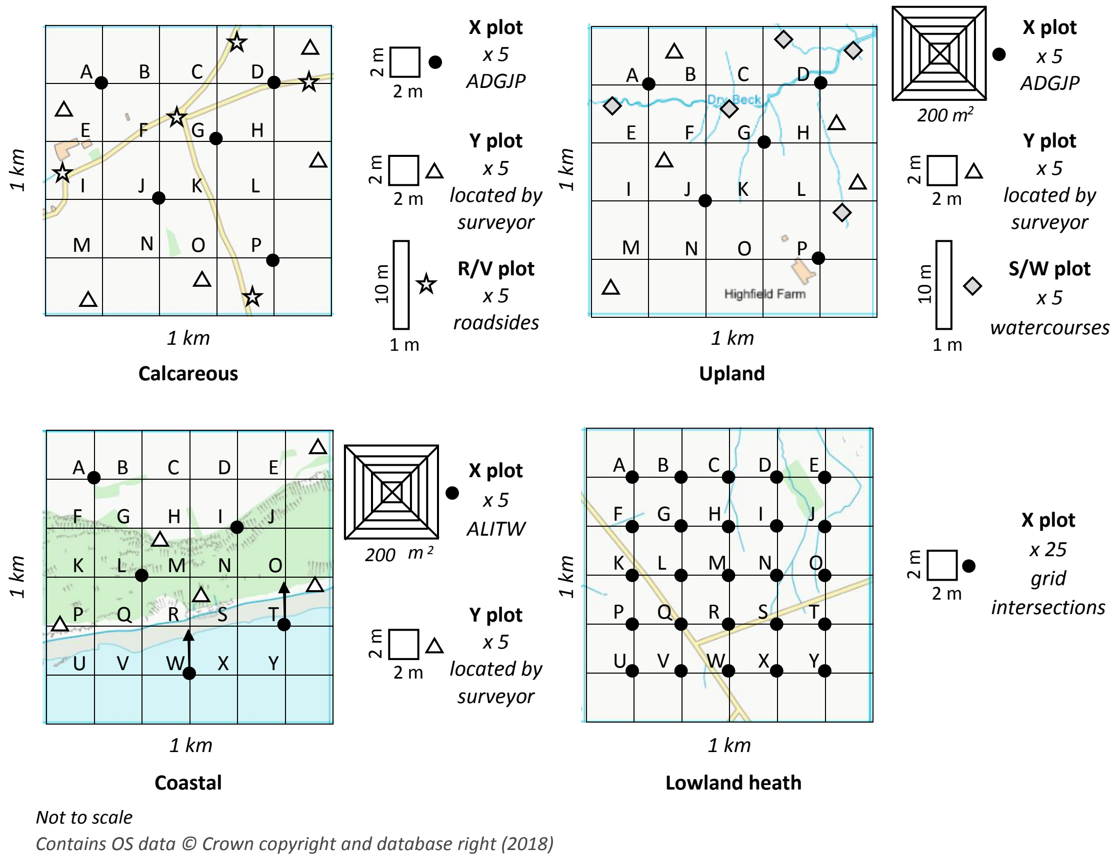


Figure 2. Gridded sampling structure for 1 km survey squares. Vegetation plots were recorded as shown, with land cover and boundary features being recorded at every grid intersection point in each square.

6 Analysis to date: key findings

At the present time, the results of the survey have been restricted to a set of contract reports, published in 1996 (Barr, 1996a, b, c, d). The previous unavailability of the data has so far resulted in limited use of the datasets, although one example has been the incorporation of the plot data in the niche models included in the Multimove package (Henrys et al., 2015), which enables users to make predictions of species occurrence from specified environmental data and allows the plotting of relationships between the occurrence of species and individual environmental covariates. A summary of the key findings reported in the 1996 reports is described in the following sections; however, the potential for further analyses is high.

6.1 Summary of results in terms of broad habitat extents

Following the Key Habitat Survey, results of stock estimates (extents) were presented in terms of land cover classes, based on those used in CS 1990 (Barr et al., 1993). Methods of classifying land cover types have since evolved (e.g. Wyatt et al., 1994). It is now possible to present estimates of habitats present in each landscape in terms of standard UK broad habitats (Jackson, 2000) and, in some cases, priority habitats. The data also offer the potential for additional work in terms of exploring priority habitats in more detail. The recorded field codes and the original land cover classes can be translated to broad habitat categories using the information presented in Table S1. Table 7 gives a summary of the broad habitat area extents (with additional coastal habitats defined in Hornung et al., 1997) provided by the Key Habitat Survey. For the purposes of comparison, the table also includes estimates for the whole of England from the national Countryside Survey (Carey et al., 2008).

In the lowland heath, calcareous grassland and coastal landscapes, only a small proportion of the landscape masks were estimated to be habitats characteristic of the landscape type (figures shown in bold in Table 7). For lowland heath: 5.2 % (dwarf shrub heath); calcareous: 1.6 % (calcareous grassland) and coastal: 11.6 % (supra-littoral rock, supra-littoral sediment, littoral sediment). The large proportion of the upland landscape which comprises characteristic habitats (56.5 %, acid grassland/bracken; dwarf shrub heath; fen, marsh and swamp; bog) reflects the less intensive use of the

Table 7. Estimates of broad habitat extents for each landscape type from the Key Habitat Survey and, for comparative purposes, in England as a whole from the Countryside Survey. Broad habitats characteristic of a landscape type are given in bold. Note that categories in italics are customised definitions used within the project rather than standard broad or priority habitats (those in bold are characteristic of a landscape).

Broad habitat (BH)		England[a]		Lowland heath			Calcareous			Coastal			Upland		
		Area ('000 ha)	%	Area ('000 ha)	% of mask	% of BH in Eng	Area ('000 ha)	% of mask	% of BH in Eng	Area ('000 ha)	% of mask	% of BH in Eng	Area ('000 ha)	% of mask	% of BH in Eng
Broadleaved, mixed and yew woodland/ coniferous woodland		1238	9.3	172	20.1	13.9	295	11.1	23.8	37	5	3.0	168	10.8	13.6
Arable and horticulture		4002	30.4	234	27.4	5.8	882	33.2	22.0	190	25.9	4.7	22	1.4	0.5
Neutral/improved grassland		4309	32.7	257	30.1	6.0	812	30.6	18.8	196	26.7	4.5	439	28.1	10.2
Calcareous grassland		30	0.2	0	0	0.0	**43**	**1.6**	**143.3**	14	1.9	46.7	0	0	0.0
Acid grassland/bracken		487	3.7	15	1.8	3.1	178	6.7	36.5	0	0	0.0	**421**	**27**	**86.4**
Dwarf shrub heath		331	2.5	**44**	**5.2**	**13.3**	50	1.9	15.1	0	0	0.0	**279**	**17.9**	**84.3**
Fen, marsh and swamp		117	0.9	0	0	0.0	16	0.6	13.6	9	1.2	7.7	**43**	**2.7**	**36.8**
Bog		140	1.1	5	0.6	3.6	32	1.2	22.8	0	0	0.0	**139**	**8.9**	**99.3**
Built-up areas and gardens		1038	7.9	108	12.7	10.4	274	10.3	26.4	200	27.2	19.3	28	1.8	2.7
Other land[b]										4	0.5	0.3	23	1.5	1.6
Littoral sediment/supra-littoral sediment	*bare shore*	1488	11.3	18	2.1	1.3	74	2.8	5.0	***26***	***3.6***		0	0	0
Littoral sediment	*salt marsh*									***37***	***5***	***100***	0	0	0
Supra-littoral rock/sediment	*maritime vegetation*									***22***	***3***		0	0	0
Total		13 180	100	854	100	–	2656	100	–	734	100	–	1562	100	–
% of Eng. in mask		100		6.5			20.1			5.6			11.8		

[a] Figures from the Countryside Survey (Centre for Ecology and Hydrology, 2009).
[b] Includes unsurveyed urban land, rivers and streams, standing open waters and canals, boundary and linear features, and coastal habitats where not otherwise specified.

Table 8. Summary of boundaries by landscape type as a proportion of the total (+ denotes present at < 1 %).

	Lowland heath	Calcareous	Coastal	Upland
Percentage of points without boundaries	32	32	55	38
Percentage of points with boundaries	68	68	45	63
Bank	4	1	10	+
Ditch	7	0	0	0
Fence	43	43	42	33
Fence and bank	2	1	3	1
Hedge	20	17	11	2
Hedge and bank	6	2	4	1
Hedge and fence	12	19	11	4
Hedge, fence and bank	5	2	3	1
Hedge and wall	0	+	1	+
Hedge, wall and fence	0	+	+	+
Wall	1	7	10	36
Wall and bank	0	+	+	+
Wall and fence	1	8	4	23
Wall, fence and bank	0	+	+	0

uplands and the extensive nature of many of the upland habitats.

More than a half of the total areas of the calcareous grassland, lowland heath and coastal landscape masks were under arable crops or managed grassland (arable and horticulture, improved/neutral grassland), reflecting the predominantly lowland distribution of these landscapes and previous intensification of agriculture (for example, Chamberlain et al., 2000). In contrast to the other landscapes, only a small proportion of the upland landscape area was classed as arable and horticulture (1.4 %), with a large proportion of the land cover consisting of semi-natural vegetation; crops were only recorded in the marginal uplands. The largest area of urban broad habitat was found in the coastal landscape (27.2 %) showing the extent of urban development in the coastal zone. The largest area of woodland (broadleaved, mixed and yew/coniferous woodland) occurred in the lowland heath mask (20.1 %) and the smallest in the coastal mask (5 %).

Figures from the Countryside Survey enable an assessment of the amount of each broad habitat within each landscape covered by the Key Habitat Survey compared with national figures for the whole of England. In the case of dwarf shrub heath, Countryside Survey estimates a stock of 331 000 ha in England. The survey of dwarf shrub heath in the lowland heathland (44 000 ha) and upland landscapes (279 000 ha) in the Key Habitat Survey gives a lower overall estimate than CS, at 323 000 ha, indicating that perhaps some small areas of heath were missed during the Key Habitat Survey. The upland habitats (incorporating acid grassland, bracken, dwarf shrub heath and bog) are covered well by the Key Habitat Survey, covering 84.3–99.3 % of the total England areas. A total of 36.8 % of the fen, marsh and swamp habitat was found in the upland areas (but is also present in lowland areas). In terms of the calcareous grassland landscape, the Key Habitat Survey estimates a total of 43 000 ha in comparison with a CS total of 30 000 ha. This perhaps confirms the fact that CS is not designed to effectively monitor or survey less common habitats such as this (Morton et al., 2011).

In the original survey reports, analysis indicated that, overall, the vegetation of the coastal landscape was the most sensitive to the changes considered (such as arable intensification, urban development, climate change, and recreation pressure). In all four landscapes, it was found that the majority of high-quality habitats were located within protected areas, potentially demonstrating the effectiveness of designation in restricting habitat loss (Hornung et al., 1997).

6.2 Summary of boundary results

The proportion of different boundary types recorded in each of the landscape masks is shown in Table 8, including the proportion of points for which there was (or was not) a boundary within 100 m. In calcareous, coastal and lowland heath landscapes, fences are the most frequent boundary type, accounting for 42–43 % of all boundaries. In the uplands, fences accounted for 33 % of all boundaries, whereas walls formed 36 %. Combinations of walls and fences accounted for a further 23 %.

Field boundaries were most common in the calcareous and lowland heath landscape areas, with 68 % of points having a boundary within 100 m, reflecting field size, cropping practices and the presence of urban features (including roads).

In coastal land, only 45 % of all grid points had a boundary within 100 m. Squares in protected, designated land had a

Table 9. Mean number of species in each habitat indicator group per plot in each landscape type in the Key Habitat Survey, with an indication of broad (BH) and priority (PH) habitats where appropriate. Habitat indicator groups characteristic of a landscape type are given in bold.

	Lowland heath		Calcareous (4 m^2 X plot)		Calcareous (4 m^2 Y plot)		Coastal (4 m^2)		Coastal (200 m^2)		Upland (4 m^2)		Upland (200 m^2)	
Habitat indicator groups	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Acid grassland/moorland species (BH acid grassland, dwarf shrub heath)	2.6	27	2.2	15	2	11	1	8	2	9	**3.9**	**23**	**6.6**	**29**
Aquatic margin species	–	–	–	–	–	–	0.4	3	0.2	1	–	–	–	–
Base-rich grassland/flush species	–	–	1	6	1.6	8	–	–	–	–	0.9	5	0.5	2
Bog/acid flush species (BH bog, BH fen, marsh and swamp)	–	–	–	–	–	–	–	–	–	–	**1.8**	**10**	**1.9**	**8**
Calcareous grassland species (BH calcareous grassland)	–	–	**0.4**	**3**	**0.6**	**3**	1.2	9	1.3	6	–	–	–	–
Damp grassland/tall herb species	–	–	0.5	3	0.8	4	0.5	4	0.8	3	–	–	–	–
Heath generalist species (PH lowland heath)	**4**	**42**	–	–	–	–	–	–	–	–	–	–	–	–
Heath specialist species (PH lowland heath)	**0.6**	**6**	–	–	–	–	–	–	–	–	–	–	–	–
Maritime species (BH supra-littoral rock, BH supra-littoral sediment, BH littoral sediment)	–	–	0	0	0	0	**2**	**15**	2.1	9	–	–	–	–
Marsh and aquatic species	–	–	0.1	1	0.8	4	–	–	–	–	–	–	–	–
Neutral/improved grassland species	–	–	–	–	–	–	–	–	–	–	4.6	27	6.3	27
Neutral grassland species	0.6	6	6.6	45	7	38	4.7	35	9.9	43	–	–	–	–
Streamside/marsh species	–	–	–	–	–	–	–	–	–	–	1.7	10	1.1	5
Upland grass species (BH acid grassland)	–	–	–	–	–	–	–	–	–	–	**2.4**	**14**	**3.9**	**17**
Weeds/alien species	0.2	2	1.7	11	2.6	14	2.1	16	4	17	0.4	2	1	4
Woodland/scrub species	1.5	16	1.4	9	1.7	9	0.6	5	1.5	6	1.4	8	1.9	8
Woodland edge/scrub species	–	–	0.9	6	1.5	8	0.6	5	1.4	6	–	–	–	–
Totals	9.5	100	14.8	100	18.6	100	13.1	100	23.2	100	17.1	100	23.2	100

lower proportion of field boundaries, indicating the greater areas of unenclosed parcels on protected land.

In the uplands, 63 % of all grid points had a boundary within 100 m. There was a clear difference between strata in the number of boundaries. Additional analyses showed the squares in the true uplands had a lower proportion of field boundaries, showing the greater areas of unenclosed land (heath and woodland) (Barr, 1996c). In designated land, and the non-designated marginal land, walls (with or without fences) formed the most frequent boundary type, followed by fences, but, in the non-designated true upland land, walls were less common and fences formed the predominant boundary type. Only 7 % of boundaries in the uplands included hedges.

6.3 Summary of vegetation plot results

The range of vegetation present is described using a classification of plants derived from statistical clustering of the species scores from DECORANA (Hill, 1979) axes into "habitat indicator groups", later developed as the Countryside Vegetation System, as described by Bunce et al. (1999). This term was coined in conjunction with the Department of the Environment, and their occurrence helps to interpret the ecological characteristics of the landscapes. The mean number of species in each of these habitat indicator groups per plot for each landscape type is shown in Table 9, along with the proportion of species in each indicator group in comparison with the total. An indication of the current broad (BH) or priority habitat (PH) to which the habitat indicator group equates is given in Table 9. Although the proportion of species from each indicator group falling into each landscape type in many cases reflects the overall extent of that type (figures in bold in Table 9), it also reflects the extent of fragmentation of some vegetation types, thus giving an indication of the quality of that type. The characteristic vegetation types were well represented in the main plots in the uplands, showing that they occur as relatively large areas. The uplands were dominated by moorland (23–29 %), bog (8–10 %), and upland grassland (14–17 %) species, but also include a variety of more lowland indicator groups, such as neutral and improved grassland species (27 %) as well as woodland species (8 %).

In calcareous landscapes, the proportion of species from the calcareous grassland habitat indicator group was only 3 % of the total. This indicates the scarcity and largely fragmented distribution of unimproved calcareous grassland even in areas with suitable geology. The proportion of species was far higher in the neutral grassland group (38–45 %) and even the acid/moorland group (11–15 %).

The habitat indicator groups with the highest proportion of species in the lowland heath landscapes were heath generalist species (42 %) and acid or moorland species (27 %). Woodland species were also well represented (16 %).

In coastal landscapes, 35–43 % of the species fell into the neutral grassland species group, followed by weeds/alien species (16–17 %). Maritime species only accounted for 9–15 % of the total.

Additional analysis showed that distribution of characteristic vegetation types demonstrated differences between designated and non-designated areas, suggesting that larger areas of characteristic vegetation occurred in the designated sample squares (Hornung et al., 1997).

7 Data availability

The datasets have been assigned digital object identifiers and users of the data must reference the data as Barr et al. (2017) and Bunce et al. (2017).

The datasets are available from the CEH Environmental Information Data Centre Catalogue (https://eip.ceh.ac.uk/data). Datasets are provided under the terms of the Open Government Licence (http://eidchub.ceh.ac.uk/administration-folder/tools/ceh-standard-licence-texts/ceh-open-government-licence/plain, http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/). The metadata are stored in the ISO 19115 (2003) schema (International Organization for Standardization, 2015) in the UK Gemini 2.1 profile (UK GEMINI, 2015). Users of the datasets will find the following documents useful (supplied as supporting documentation with the datasets from a link on the DOI landing page): Barr (1992, 1993).

8 Conclusion

During recent decades there has been increasing concern over the loss of a number of valued landscapes and their associated characteristic habitats. A number of policies have been introduced to protect and enhance the remaining areas of these characteristic habitats. The UK biodiversity action plan (and the EU Habitats Directive) has also set targets for the protection of threatened species and habitats. However, overall, there is inadequate information with which to judge the status and quality of these and how they are changing at a national level. Together, the land cover and vegetation data described in the present paper provide an important baseline offering the potential for the monitoring and evaluation of threats to the landscapes and their characteristic habitats. The data also offer information useful for evaluating the quality and ecological characteristics of the surveyed landscape types in relation to a range of potential drivers.

It seems likely that further declines in ecological quality may have occurred since the survey bearing in mind current trends, but the extent of these could only be determined by a monitoring programme, for which this survey provides a useful framework. The Countryside Survey has demonstrated the robustness of a similar database for such a repeat.

According to the findings to date, it could be expected that changes are more likely in unprotected, undesignated land in the uplands than in protected, designated land in coastal, heath and calcareous grasslands. In general, previous analysis of these data has shown that the areas protected by legislation (designated) are of higher ecological quality than those in non-designated areas. This result could indicate that such designations may therefore provide protection for threatened habitats but it may also reflect the original designation of high-quality habitats. This is valuable information in the targeting of initiatives and funding designed to restore the given habitats.

The datasets provide a broadly defined distribution in England of four landscapes of interest including the broad habitats characteristic of the landscapes, as well as areas with potential for these habitats. These data form valuable contextual information for further specific surveys and monitoring. The datasets also provide an objective characterisation and quantification of the land cover and vegetation within the defined areas of these landscapes by field survey of a stratified random sample of 1 km squares within each landscape. The resultant data have been used to assess the distribution of species representative of the characteristic habitats and in the different sampling strata of the landscapes, and they offer much potential for further work.

The survey was the first time that a statistically rigorous assessment of ecological quality has been attempted across such a wide range of ecologically important habitats using similar methods and standardised protocols at a national level. The standardised design of the survey offers the opportunity for the possible integration with future monitoring surveys of the status of the British countryside, as an element of the Countryside Survey programme. The additional targeted 1 km sampling squares of the Key Habitat Survey could be surveyed as an additional element within the Countryside Survey field survey to add value and yield additional information regarding the targeted landscapes in question, should resources allow. The location of the vegetation plots have been permanently marked to facilitate future resurvey and are thus able to be monitored over time and, as stated above, would facilitate long-term ecological monitoring linked to a range of drivers. Consideration should be given to the inclusion of these additional targeted sites in the next full Countryside Survey in Britain, for which an addition to the series is now overdue (the latest updates may be found at www.countrysidesurvey.org.uk).

The Supplement related to this article is available online at https://doi.org/10.5194/essd-10-899-2018-supplement.

Author contributions. CMW prepared the manuscript with contributions from all co-authors and is the current database manager for the Land Use Research Group at CEH Lancaster. The sampling framework and survey strategy was based on methods designed by RGHB, and the field survey was overseen by CJB.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank all the land owners who kindly gave permission to survey their holdings in the survey sample squares in 1992 and 1993. Without their co-operation and assistance the Countryside Survey and related surveys would not exist. We also acknowledge and thank all the field surveyors involved in each field campaign (Henry Adams, Tanya Barden, Elizabeth Biron, Roger Cummins, John Davis, John Day, Richard Hewison, Gabby Levine, Amanda Marler, Elizabeth McDonnell, Karen Pollock, Sam Walters, Mike Webb). We thank Caroline Hallam for data management during the project in the 1990s. The survey was funded by the Department of the Environment (DOE) (now the Department for Environment, Food and Rural Affairs (DEFRA). We thank two anonymous reviewers whose comments have improved the manuscript considerably.

Edited by: David Carlson
Reviewed by: two anonymous referees

References

Aiba, S.-i. and Kitayama, K.: Structure, composition and species diversity in an altitude-substrate matrix of rain forest tree communities on Mount Kinabalu, Borneo, Plant Ecol., 140, 139–157, https://doi.org/10.1023/A:1009710618040, 1999.

Barr, C. J.: Countryside Survey 1990 field handbook, Institute of Terrestrial Ecology, Grange-over-Sands, 1990.

Barr, C. J.: Changes in Key Habitat: A Survey of Lowland Heath, Field Survey Handbook, ITE Merlewood, Grange over Sands, 1992.

Barr, C. J.: Changes in Key Habitat: Surveys of Chalk & Limestone Grassland, Coastal, and Upland Landscapes, Field Handbook. Draft 3, Institute of Terrestrial Ecology, Grange-over-Sands, 1993.

Barr, C. J.: Current status and prospects for key habitats in England. Part 1: Lowland heath landscapes, Department of the Environment, Transport and the Regions, Institute of Terrestrial Ecology, Grange over Sands, 1996a.

Barr, C. J.: Current status and prospects for key habitats in England. Part 2: Calcareous grassland landscapes, Department of the Environment, Transport and the Regions, Institute of Terrestrial Ecology, Grange over Sands, 1996b.

Barr, C. J.: Current status and prospects for key habitats in England. Part 3: Upland landscapes, Department of the Environment, Transport and the Regions, Institute of Terrestrial Ecology, Grange over Sands, 1996c.

Barr, C. J.: Current status and prospects for key habitats in England. Part 4: Coastal landscapes, Department of the Environment, Transport and the Regions, Institute of Terrestrial Ecology, Grange over Sands, 1996d.

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Fuller, R. M., Furse, M. T., Gillespie, M. K., Groom, G. B., Hallam, C. J., Hornung, M., Howard, D. C., and Ness, M. J.: Countryside Survey 1990: main report. (Countryside 1990 vol. 2), Department of the Environment, London, UK, 1993.

Barr, C. J., Bunce, R. G. H., Cummins, R. P., Hallam, C. J., Hornung, M., and Wood, C. M.: Habitat and vegetation data from an ecological survey of terrestrial key habitats in England, 1992–1993, NERC Environmental Information Data Centre, https://doi.org/10.5285/7aefe6aa-0760-4b6d-9473-fad8b960abd4, 2017.

Bridgewater, S., Ibanez, A., Ratter, J., and Furley, P.: Vegetation classification and floristics of the savannas and associated wetlands of the Rio Bravo Conservation and Management Area, Belize, Edinburgh Journal of Botany, 59, 421–442, 2002.

British Geological Survey: 1:625 000 Geology: Great Britain and Northern Ireland, available at: http://www.bgs.ac.uk/products/digitalmaps/digmapgb_625.html, last access: 8 October 2017.

Broyer, J. and Curtet, L.: The National Hay-meadow Ecosystem Observatory: a tool to monitor and understand the ongoing changes in France, Integrating Efficient Grassland Farming and Biodiversity – Proceedings of the 13th International Occasional Symposium of the European Grassland Federation Tartu, Estonia, 29–31 August 2005, Grassland Science in Europe, 10, 52 pp., 2005.

Bunce, R. G. H. and Shaw, M. W.: A standardised method for ecological survey, J. Environ. Manage., 1, 239–258, 1973.

Bunce, R. G. H. and Smith, R. S.: An ecological survey of Cumbria, Structure Plan Working Paper 4, Cumbria County Council and Lake District Special Planning Board, Kendal, UK, 1978.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE Land Classification of Great Britain 1990, NERC Environmental Information

Data Centre, https://doi.org/10.5285/ab320e08-faf5-48e1-9ec9-77a213d2907f, 1990.

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D. C., and Lane, A. M. J.: ITE Merlewood Land Classification of Great Britain, J. Biogeogr., 23, 625–634, https://doi.org/10.1111/j.1365-2699.1996.tb00023.x, 1996a.

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., and Howard, D. C.: The ITE Land Classification: providing an environmental stratification of Great Britain, Environ. Monit. Assess., 39, 39–46, https://doi.org/10.1007/978-94-009-1653-1_5, 1996b.

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., Howard, D. C., Scott, W. A., Smart, S. M., Van de Poll, H. M., and Watkins, J. W.: Vegetation of the British countryside – the Countryside Vegetation System, Grange-over-Sands Institute of Terrestrial Ecology, ECOFACT Volume 1, 225 pp., 1999.

Bunce, R. G. H., Parr, T. W., Ullyett, J., Hornung, M., Gerard, F., Bull, R., Cox, R., and Brown, N. J.: Spatial masks for calcareous, coastal, upland and lowland heath landscapes in England [Key Habitats 1992-93], NERC Environmental Information Data Centre, https://doi.org/10.5285/dc583be3-3649-4df6-b67e-b0f40b4ec895, 2017.

Carey, P. D., Wallis, S., Chamberlain, P. M., Cooper, A., Emmett, B. A., Maskell, L. C., McCann, T., Murphy, J., Norton, L. R., Reynolds, B., Scott, W. A., Simpson, I. C., Smart, S. M., and Ullyett, J. M.: Countryside Survey: UK Results from 2007, NERC/Centre for Ecology & Hydrology, Lancaster, 2008.

Centre for Ecology and Hydrology: Countryside Survey: England Results from 2007 NERC/Centre for Ecology & Hydrology, Lancaster, CEH Project Number: C03259, 119 pp., 2009.

Chamberlain, D. E., Fuller, R. J., Bunce, R. G. H., Duckworth, J. C., and Shrubb, M.: Changes in the abundance of farmland birds in relation to the timing of agricultural intensification in England and Wales, J. Appl. Ecol., 37, 771–788, https://doi.org/10.1046/j.1365-2664.2000.00548.x, 2000.

Countryside Stewardship: https://www.gov.uk/government/collections/countryside, last access: 21 August 2017.

Cranfield University: National Soil Map of England and Wales – NATMAP, available at: http://www.landis.org.uk/data/natmap.cfm, last access: 8 October 2017.

Dargie, T.: Sand Dune Survey of Great Britain – A National Inventory. Part 2 Scotland, Joint Nature Conservation Committee, Coastal Conservation Branch, Peterborough, UK, 1993.

Dargie, T.: Sand Dune Survey of Great Britain – A National Inventory. Part 3 Wales, Joint Nature Conservation Committee, Coastal Conservation Branch, Peterborough, UK, 1995.

Department of the Environment: Biodiversity: The UK Steering Group Report: Volume 2: Action Plans, HMSO, London, UK, 1995.

Dolman, P. M. and Land, R.: Lowland heathland 10, in: Managing habitats for conservation, edited by: Sutherland, W. J. and Hill, D. A., Cambridge University Press, 267–291, 1995.

Doody, P. and Rooney, P.: Special issue-conservation and management of sea cliffs, J. Coast. Conserv., 19, 757–760, 2015.

Dramstad, W. E., Fjellstad, W., Strand, G.-H., Mathiesen, H. F., Engan, G., and Stokland, J. N.: Development and implementation of the Norwegian monitoring programme for agricultural landscapes, J. Environ. Manage., 64, 49–63, 2002.

Emmett, B. E. and GMEP team: Glastir Monitoring & Evaluation Programme, Final Report to Welsh Government – Executive Summary (Contract reference: C147/2010/11), NERC/Centre for Ecology & Hydrology, Bangor, UK, 2017.

EuMon: EU-wide monitoring methods and systems of surveillance for species and habitats of Community interest, available at: http://eumon.ckff.si/index1.php, last access: 8 August 2017.

exegesis SDM Ltd. and Doody, J. P.: Development of a coastal vegetated shingle inventory for England, available at: http://publications.naturalengland.org.uk/publication/46007?category=43007 (last access: 9 May 2018), 2009.

Farrell, L.: The different types and importance of British heaths, Bot. J. Linn. Soc., 101, 291–299, 1989.

Fuller, R. M.: The changing extent and conservation interest of lowland grasslands in England and Wales: A review of grassland surveys 1930–1984, Biol. Conserv., 40, 281–300, https://doi.org/10.1016/0006-3207(87)90121-2, 1987.

Fuller, R. M., Groom, G. B., Jones, A. R., and Thomson, A. G.: Land Cover Map 1990 (25 m raster, GB), NERC Environmental Information Data Centre, https://doi.org/10.5285/3d974cbe-743d-41da-a2e1-f28753f13d1e, 1993.

Gibson, C. and Brown, V.: The nature and rate of development of calcareous grassland in southern Britain, Biol. Conserv., 58, 297–316, https://doi.org/10.1016/0006-3207(91)90097-S, 1991.

Gillis, M., Omule, A., and Brierley, T.: Monitoring Canada's forests: the national forest inventory, Forest. Chron., 81, 214–221, https://doi.org/10.5558/tfc81214-2, 2005.

Government Digital Service: Find protected areas of countryside: https://www.gov.uk/check-your-business-protected-area, last access: 4 October 2018.

Henrys, P. A., Butler, A., Jarvis, S., Smart, S. M., and Fang, Z.: MultiMOVE Model: Ensemble niche modelling of British vegetation v2.0.1, NERC Environmental Information Data Centre, https://doi.org/10.5285/94ae1a5a-2a28-4315-8d4b-35ae964fc3b9, 2015.

Hill, M. O.: DECORANA – A FORTRAN program for detrended correspondence analysis and reciprocal averaging, Cornell University, USA, 1979.

Hintermann, U., Weber, D., Zangger, A., and Schmill, J.: Biodiversity Monitoring in Switzerland, BDM, Interim Report, Swiss Agency for the Environment, Berne, Switzerland, Forests and Landscape, Environmental Series No. 342, p. 20, 2002.

Hope-Simpson, J.: On the errors in the ordinary use of subjective frequency estimations in grassland, J. Ecol., 28, 193–209, https://doi.org/10.2307/2256169, 1940.

Hornung, M. F., Barr, C. J., and Bunce, R. G. H.: Current status and prospects for key habitats in England. Part 6: Summary report, Institute of Terrestrial Ecology, Grange-over-Sands, 1997.

International Organization for Standardization: ISO 19115:2003: http://www.iso.org/iso/catalogue_detail.htm?csnumber=26020, last access: 10 December 2015.

Jackson, D.: Guidance on the Interpretation of the Biodiversity Broad Habitat Classification (Terrestrial and Freshwater Types): Definitions and the Relationship with Other Habitat Classifications, JNCC, Report No. 307, 73 pp, available at: http://jncc.defra.gov.uk/page-2433 (last access: 9 May 2018), 2000.

Jerram, R., Clayden, D., and Rees, S.: North York Moors National Park Upland Vegeattion Survey, English Nature, Kendal, UK, 1998.

Joint Nature Conservation Committee: Designations associated with status based lists, http://jncc.defra.gov.uk/page-3425, last access: 12 April 2018.

Keymer, R. and Leach, S.: Calcareous grassland – a limited resource in Britain, Calcareous grasslands: ecology and management, 11–20, 1990.

Mackney, D., Hodgson, J., Hollis, J., and Staines, S.: Legend for the 1 : 250, 000 soil map of England and Wales, Soil Survey of England and Wales, Rothamsted, Harpenden, UK, 1983.

Maddock, A.: UK Biodiversity Action Plan; Priority Habitat Descriptions, available at: http://jncc.defra.gov.uk/PDF/UKBAP_PriorityHabitatDesc-Rev2011.pdf (last access: 9 May 2018), 2008.

Marrs, R., Hicks, M., and Fuller, R.: Losses of lowland heath through succession at four sites in Breckland, East Anglia, England, Biol. Conserv., 36, 19–38, 1986.

Maskell, L. C., Norton, L. R., Smart, S. M., Carey, P. D., Murphy, J., Chamberlain, P. M., Wood, C. M., Bunce, R. G. H., and Barr, C. J.: Countryside Survey. Field Mapping Handbook, Centre for Ecology & Hydrology, Lancaster, CS Technical Report No. 1/07, 143 pp., 2008a.

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., and Barr, C. J.: Vegetation Plots Handbook, Centre for Ecology and Hydrology, Lancaster, CS Technical Report No. 2/07, 64 pp., 2008b.

Moore, N. W.: The Heaths of Dorset and their Conservation, J. Ecol., 50, 369–391, https://doi.org/10.2307/2257449, 1962.

Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R., and Simpson, I.: Final Report for LCM2007 – the new UK land cover map, NERC/Centre for Ecology & Hydrology, WallingfordCS Technical Report No. 9/07, 115 pp., 2011.

Natural England: Digital Boundary Data for Designated Wildlife Sites and related information, http://www.gis.naturalengland.org.uk/pubs/gis/tech_ds.htm (last access: 9 May 2018), 2017a.

Natural England: Magic datasets: http://www.natureonthemap.naturalengland.org.uk/magicmap.aspx, last access: 8 September 2017b.

Ovenden, G. N., Swash, A. R. H., and Smallshire, D.: Agri-environment schemes and their contribution to the conservation of biodiversity in England, J. Appl. Ecol., 35, 955–960, https://doi.org/10.1111/j.1365-2664.1998.tb00014.x, 1998.

Poschlod, P. and WallisDeVries, M. F.: The historical and socioeconomic perspective of calcareous grasslands—lessons from the distant and recent past, Biol. Conserv., 104, 361–376, https://doi.org/10.1016/S0006-3207(01)00201-4, 2002.

Prosser, M. V. and Wallace, H. L.: Countryside Survey 1990 Quality Assurance Exercise. Report to Institute of Terrestrial Ecology, Merlewood Research Station, Canon Pyon, Herefordshire, UK, 1992.

Prosser, M. V. and Wallace, H. L.: Quality Assurance Exercise, Report to Centre of Ecology & Hydrology, Lancaster, UK, 23 pp., 2008.

Provoost, S., Ampe, C., Bonte, D., Cosyns, E., and Hoffmann, M.: Ecology, management and monitoring of grey dunes in Flanders, J. Coast. Conserv., 10, 33–42, https://doi.org/10.1652/1400-0350(2004)010[0033:EMAMOG]2.0.CO;2, 2004.

Pywell, R. F., Putwain, P. D., and Webb, N. R.: The decline of heathland seed populations following the conversion to agriculture, J. Appl. Ecol., 34, 949–960, https://doi.org/10.2307/2405285, 1997.

Radley, G. and Dargie, T.: Sand dune vegetation survey of Great Britain: a national inventory, Martin Walters, Joint Nature Conservation Committee, Peterborough, 392 pp., 1994.

Rodwell, J. S.: National vegetation classification: Users' handbook, Joint Nature Conservation Committee, Peterborough, UK, 2006.

Romão, C.: Interpretation manual of European Union habitats, version EUR 28, available at: http://ec.europa.eu/environment/nature/legislation/habitatsdirective/docs/Int_Manual_EU28.pdf (last access: 9 May 2018), 2013.

Rose, R., Webb, N., Clarke, R., and Traynor, C.: Changes on the heathlands in Dorset, England, between 1987 and 1996, Biol. Conserv., 93, 117–125, https://doi.org/10.1016/S0006-3207(99)00047-6, 2000.

Ryle, T., Connolly, K., Murray, A., and Swan, M.: Coastal Monitoring Project 2004–2006: A report prepared for the National Parks and Wildlife Service, Research Branch, Peterborough, UK, 2007.

Sneddon, P., Randall, R., and Joint Nature Conservation Committee, P.: Coastal vegetated shingle structures of Great Britain: Appendix 2. Shingle sites in Scotland, JNCC, Cambridge, UK, 1994.

Špulerová, J.: Vegetation monitoring of Mutne peat bog, Slovakia, in: Monitorovanie a hodnotenie stavu životného prostredia VIII, Technická univerzita vo Zvolene, Zvolen, Slovak Republic, 2009.

Ståhl, G., Allard, A., Esseen, P.-A., Glimskär, A., Ringvall, A., Svensson, J., Sundquist, S., Christensen, P., Torell, Å. G., and Högström, M.: National Inventory of Landscapes in Sweden (NILS) – scope, design, and experiences from establishing a multiscale biodiversity monitoring system, Environ. Monit. Assess., 173, 579–595, https://doi.org/10.1007/s10661-010-1406-7, 2011.

Stevens, D., Blackstock, T., Smith, S., and Bosanquet, S.: Lowland grassland survey of Wales, British Wildlife, 18, 314–323, 2007.

Stewart, A., Pearman, D., and Preston, C.: Scarce plants in Britain, JNCC, Peterborough, UK, 1994.

Tallis, J.: Forest and Moorland in the South Pennine Uplands in the Mid-Flandrian Period: III. The Spread of Moorland–Local, Regional and National, J. Ecol., 79, 401–415, 1991.

Thompson, D. B. A., MacDonald, A. J., Marsden, J. H., and Galbraith, C. A.: Upland heather moorland in Great Britain: A review of international importance, vegetation change and some objectives for nature conservation, Biol. Conserv., 71, 163–178, https://doi.org/10.1016/0006-3207(94)00043-P, 1995.

UK Biodiversity Steering Group: Biodiversity: The UK Steering Group Report: Volume 2: Action Plans, HMSO, London, UK, 1995.

UK GEMINI: https://www.agi.org.uk/agi-groups/standards-committee/uk-gemini, last access: 9 May 2018.

United States Forest Service: National Forest Health Monitoring Program: http://fhm.fs.fed.us/fact/index.htm, last access: 15 May 2015.

Usher, M. and Thompson, D.: Variation in the upland heathlands of Great Britain: conservation importance, Biol. Conserv., 66, 69–81, https://doi.org/10.1016/0006-3207(93)90136-O, 1993.

Webb, N.: Heathlands, Collins, London, UK, 1986.

Webb, N. and Haskins, L.: An ecological survey of heathlands in the Poole Basin, Dorset, England, in 1978, Biol. Conserv., 17, 281–296, https://doi.org/10.1016/0006-3207(80)90028-2, 1980.

Wilson, P., Wheeler, B., Reed, M., and Strange, A.: A survey of selected agri-environment grassland and heathland creation and restoration sites: Part 2, Natural England Commissioned Reports, Number 107, 95 pp., 2013.

Wiser, S. K., Bellingham, P. J., and Burrows, L. E.: Managing biodiversity information: development of New Zealand's National Vegetation Survey databank, New Zeal. J. Ecol., 25, 1-17, 2001.

Wood, C. M. and Bunce, R. G. H.: Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring, Earth Syst. Sci. Data, 8, 89–103, https://doi.org/10.5194/essd-8-89-2016, 2016.

Wood, C. M., Smart, S. M., and Bunce, R. G. H.: Woodland Survey of Great Britain 1971–2001, Earth Syst. Sci. Data, 7, 203–214, https://doi.org/10.5194/essd-7-203-2015, 2015.

Wood, C. M., Smart, S. M., Bunce, R. G. H., Norton, L. R., Maskell, L. C., Howard, D. C., Scott, W. A., and Henrys, P. A.: Long-term vegetation monitoring in Great Britain – the Countryside Survey 1978–2007 and beyond, Earth Syst. Sci. Data, 9, 445–459, https://doi.org/10.5194/essd-9-445-2017, 2017.

Wood, C. M., Bunce, R. G. H., Norton, L. R., Maskell, L. C., Smart, S. M., Scott, W. A., Henrys, P. A., Howard, D. C., Wright, S. M., Brown, M. J., Scott, R. J., Stuart, R. C., and Watkins, J. W.: Ecological landscape elements: long-term monitoring in Great Britain, the Countryside Survey 1978–2007 and beyond, Earth Syst. Sci. Data, 10, 745–763, https://doi.org/10.5194/essd-10-745-2018, 2018.

Wyatt, B. K., Greatorex-Davies, J. N., Hill, M. O., Parr, T. W., Bunce, R. G. H., and Fuller, R. M.: Comparison of land cover definitions, Department of the Environment, London, (Countryside 1990 series Volume 3), 1994.

sciendo
10.2478/jlecol-2018-0015 Journal of Landscape Ecology (2018), **Vol: 11 / No. 3.**

THE ECOLOGY OF BRITISH UPLAND LANDSCAPES. I. COMPOSITION OF LANDSCAPES, HABITATS, VEGETATION AND SPECIES

ROBERT G.H. BUNCE[1], CLAIRE M. WOOD[2*], SIMON M. SMART[2]

[1]*Estonian University of Life Sciences, Kreuzwaldi 5, 51014 Tartu, Estonia*
[2]*Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster, LA1 4AP, UK.*
[*]*Corresponding author e-mail: clamw@ceh.ac.uk*

Received: 30th October 2018, **Accepted:** 21th December 2018

ABSTRACT

A primary requirement for policy objectives is reliable figures on the composition of any region. Currently there is no comprehensive, definitive set of statistics for the British Uplands, hence the present paper. An overview of the background to the region is first provided, together with some examples of the available figures and a discussion of their limitations. The paper uses a formal structure, with landscapes at the highest level followed by habitats, then vegetation, and finally species, with exact definitions of the categories applied at all levels. The figures are produced from a survey of stratified, random one kilometre squares. The tables give comprehensive figures for Great Britain (GB) as a whole, and also England, Wales and Scotland.

The Uplands are shown to cover 38 % of the country. In terms of UK Broad Habitats, Bog is the most common overall (2062 k ha). It is estimated that 41 % of upland vegetation in Britain is grazed by sheep, and *Cervus elephus* (red deer) are particularly evident in Scotland. Walls (mainly drystone) are the most important linear feature (84 k km) but hedgerows (30 k km) are also widespread. The major vegetation classes are those linked to moorlands and bogs (about 25 %) but those associated with fertile soils are also common (10 %). In terms of species, *Potentilla erecta* (tormentil) is the most frequent species with four other acid grassland species in the top ten. *Calluna vulgaris* (ling heather) has the highest cover in Great Britain (14.8 %).

Keywords*:* stratified random sampling, standard habitat categories, comprehensive national estimates, vascular plant species, linear features, Brexit.

BACKGROUND

The initial comparable analysis of the British Uplands was that of Bunce (1987) using the first set of Countryside Survey (Carey *et al*., 2008) data collected in 1978 (Bunce, 1979). However, data sets collected in subsequent surveys are more comprehensive, enabling analyses (for example of vegetation data) that had not been carried out in 1987. It is therefore timely to repeat the basic analyses, but also to add the greater detail that is now available. Furthermore, as discussed below, official figures for upland habitats are not comprehensive, and moreover are not the product of robust statistical sampling and analysis. Also, as discussed by Monbiot in the British national press (Monbiot, 2017b), there is considerable public interest in the extent and composition of the Uplands because of their importance for leisure activities and biodiversity, as well as for their valued landscapes, as reflected by the extent of protected areas. Finally, the consequences of Brexit are particularly likely to have an impact in the Uplands, so it is therefore valuable to have a statement of their current composition at this time.

The first section of the paper therefore provides a definition of the region before discussing the extent of available figures. The methods of data collection are then described before the results are presented.

INTRODUCTION

After the Ice Age, forest colonised the whole of Britain, with the exception of the high mountains and exposed northern coasts, as summarised by Bunce *et al.* (2014b). Progressive deforestation then took place so that by the 19th Century, under 10 % of the original forest remained. The Victorian sports of deer stalking and grouse shooting further accentuated the process.

After the Second World War, the introduction of agricultural subsidies, followed by accession to the European Union (EU), led to a further increase in sheep numbers as well as a decline in cattle numbers. This historical process is documented by Ratcliffe and Thompson (1988) who updated the classic text of Pearsall (1950), who described the ecological characteristics of the region. Recently Reed *et al*. (2009) have provided detailed discussions of the drivers of change and a review of the dynamics of the Uplands. The particularly important drivers of fire and climate change are covered by Davies *et al.* (2016) and House *et al*. (2010) respectively. The former present a discussion of the significance of the use of fire as a management tool in the Uplands, emphasizing that fire has significant effects on biodiversity and ecosystem function. The latter paper emphasises that there is evidence that climate change is already taking place in the Uplands. Research using a climate envelope model suggests that as much as 50 % of the Uplands will be exposed to climate stress by the end of the 21st Century (Berry *et al*., 2002). However detailed interpretations of the implications of this result are more difficult. A useful description of farming in the Uplands is given by Clother and Finch (2010) with enterprises varying from traditional sheep raising to more intensive use of fertile valleys .

On account of the intense interest in the topic and the need for an update of the increasing pressures, the British Ecological Society convened a meeting to discuss the issues as reported by Evans *et al*. (2017). Many disparate topics were covered, leading to the conclusion that a transformational moment is now evident. The strength of views discussed at the meeting (Evans *et al*., 2017) was focussed on vertebrate predators, climate change and nitrogen deposition.

Paper II (Bunce *et al.*, 2018) provides an extensive discussion of the impacts of policy on the ecology of the Uplands, needing a separate text to cover the complex issues involved.

The structure of the analyses follows the principles described by Bunce (1999). The approach is and overall results of the Countryside Survey project are described by Norton *et al*. (2012) and Carey *et al*. (2008), and further details are provided at: www.countrysidesurvey.org.uk. Each level is explicit and is formed of complexes of classes of the previous stage. The levels are as follows:

1. **Landscape**: The general concept of the Uplands is usually applied at the landscape level involving complexes of habitats (see below) and vegetation. There are therefore habitats usually associated with the lowlands in upland valleys, such as fertile grassland. Such habitats are not only important visually, but are often important for biodiversity, for example as feeding grounds for geese (particularly greylag, *Anser anser*). Historically there was more variation in land covers such as crops. In the present paper, Upland landscapes are determined by the appropriate classes of an environmental classification derived from statistical analysis of environmental data from all 1 km squares in Great Britain (Bunce *et al*., 1996).
2. **Habitats**: in the present paper, the Broad Habitats of the UK Biodiversity Action Plan (Jackson, 2000) are used. (This is also the level used in the EU, in the Habitats Directive (Council Directive 92) for Natura 2000 (European Commission, 1992), and the two classifications have been linked by Bunce *et al*. (2012) to enable comparisons). The scale of habitats varies according to the species involved. Thus, a butterfly such as the Adonis blue (*Lysandra bellargus*) may occupy only a patch of calcareous grassland, but a large bird may use several habitats (for example, a golden eagle (*Aquila chrysaetos*) may inhabit cliffs, trees, and a range of moorland and heathland habitats). In the present paper the habitats are recorded in the field as standardized spatially explicit units within dispersed random 1 km squares stratified according to the environmental classification (see below).
3. **Vegetation classes**: in the present paper, the vegetation classes ('Countryside Vegetation Classes') used are determined by statistical analysis of plant species recorded from dispersed random 200m square plots within the same 1 km squares (5 per square), as defined in Bunce *et al*. (1999b).
4. **Vegetation species**: these are important in their own right because they have links to other taxa, for example *Calluna* has links to grouse (*Lagopus lagopus scotica*). They are also key determinants of many habitats and form the basic structure of vegetation classes.

Definition of Uplands and extant figures

A paper on afforestation in the Uplands (Bunce *et al*., 2014b) used a comparable procedure to that given in the present paper to define the region, in that statistically derived relevant environmental classes were used to produce the map, as shown in Figure 1 below.

There is often a degree of circularity in the definitions used in the literature. Two of the best definitions are:

1. Averis et al. (2004), where upland is defined as the areas of the country which have an upland environment regardless of altitude. They are usually wetter and cooler than the lowlands and are windier and usually with poor soils.
2. Definition 1 (above) is consistent with that used to describe Less Favoured Areas (LFAs) in the EU agricultural grant system which is as follows: an area with natural handicaps (lack of water, poor climate and associated short crop season and infertile soils) (European Commission, 2018). The social criterion of declining populations as used in the EU definition of LFAs certainly applies, although it is not ecological.

The most complete figures are provided within the Biodiversity Action Plan Reporting System (BARS) (JNCC, 2012). However, not only do the figures not cover the entire landscape, as described below, but the methods of derivation are not discussed, presumably because they are extracted from a variety of sources. Areas of the following habitats were not available from this source:

1. Habitats usually associated with lowlands and not characteristic of the uplands, but are commonly present in upland landscapes, especially valleys (for example, fertile grasslands).
2. Inland rock and scree.
3. Mountain heath and willow scrub.
4. Upland flushes and swamps.

MATERIALS AND METHODS

Landscapes

In the present paper, upland landscapes are defined according to an integrated multivariate analysis (Hill & Šmilauer, 2005) of environmental data (for example climatic data, topographic data, human geographical features and geology data) originally providing 32 classes (or 'Land Classes') within Great Britain (GB). This classification was first developed in the late 1970s, and is termed the 'ITE Land Classification' (Bunce *et al*., 1996). The same classification was later modified to accommodate devolution in Scotland and Wales, leading to 45 rather than the original 32 classes by 2007 (Bunce *et al*., 2007). The interpretation of the environmental characteristics of the classes enables those that fit the definition of the Uplands to be identified, as shown in Fig. 1. All the data described below have been collected as part of GB Countryside Survey, as described by Norton *et al*. (2012), using the environmental Land Classification framework, developed over 40 years, as described by Sheail & Bunce (2003). The analytical procedure used to produce national estimates from the data collected within this framework calculates the mean values for the various parameters collected in the 1km survey squares, for each of the Land Classes. Estimates for each Land Class are then combined for either the whole of Great Britain, or for each of England, Scotland and Wales. The national estimates of Broad Habitats for 2007 (Brown *et al*., 2014a) with standard errors are publicly available via the NERC Environmental Information Data Centre along with the linear feature data (Brown *et al*., 2014b) and vegetation data (Bunce *et al*., 2014a).

Site Selection

The procedure for site selection for survey applies to both habitats and vegetation and is described below. Sites are based on a series of 1km survey units, selected on a random basis from within the Land Classification framework, described above. Initially, in the late 1970s, it was decided that eight 1km square survey sites from each of the 32 classes were the minimum to get a representative sample for the whole of Britain. Accordingly, eight dispersed random squares were drawn from each of the classes giving a total of 256 squares, of which the majority were surveyed in the summer 1978. In subsequent surveys, the sample number was increased in proportion to the size and number of the Land Classes as follows: 1984 - 382 squares, 1990 - 506 squares, 1998 – 569 squares, 2007 – 591 squares. In 2007, 238 of these squares were located in upland Land Classes.

Habitats

Information is recorded for landscape features, habitats and land cover and for each of the 1 km survey squares ensuring complete coverage (Wood *et al*., 2018). Initially the data were recorded on standard forms as described in the field handbook (Bunce, 1978). In the most recent survey in 2007, records were made digitally, using rugged field computers. Each parcel is ascribed a primary land cover code, as given in the field handbook (Maskell *et al*., 2008), followed by details of the species composition of vascular plants with over 10 % cover and information on land use or management (such as the type of grazing animals). These data enable each parcel to be allocated to the Broad and Priority Habitats of the UK Action Plan (Jackson, 2000; Maddock, 2008). Definitions and methodologies of individual vegetation and landscape mapping survey components are documented in Wood *et al*. (2017) and Wood *et al*. (2018). Full details of the Countryside Survey are given by Norton et al. (2012) and Carey *et al*. (2008). Other European habitat classifications can also be derived from these classes as shown by Bunce *et al*. (2012). Linear features such as hedgerows, and point features such as trees, are also recorded using standard codes.

Vegetation

Vegetation data were recorded from five dispersed random main plots in each 1 km square and, initially in 1978, from six plots placed along hedgerows, water courses and roadsides (later increased according to landscape variations in survey squares) (Wood *et al*., 2017). The presence and cover of all vascular plants, and a selected range of readily identifiable bryophytes and macro-lichens, were recorded on standard waterproof paper sheets, and later on hand held field computers. The main plots are 200m square and the linear plots 10 m by 1m, as described by Wood *et al*. (2017) including information about the re-location procedure. Management information was also recorded. A total of 5953 plots of all types were recorded in upland Britain in 2007, including 1136 main (X) plots (E: 252, S: 649, W: 235).

Species

The most frequent and highest cover species were derived from the main plots, and are provided for GB and England, Scotland and Wales separately.

RESULTS

Landscapes

The distribution of the upland Land Classes, derived from the ITE Land Classification (Bunce *et al*., 1996), is shown in Figure 1, and the related area figures in Table 1. The distribution patterns clearly show the moorlands of Exmoor and Dartmoor and the Malvern Hills in South West England, and the dominance of uplands in Wales. In the North of England, there are the Pennines, the Lake District and the North York Moors. In Scotland, the Southern Uplands and the Highlands are clearly uplands, but also the lower land of the North West and all the Western Isles are marginal uplands, because of the more northern climate, similar to that in South West Norway. The map corresponds closely to the appropriate classes of the climate map of Europe of Metzger *et al.* (2005).

Table 1 shows the contrast between the three countries, with England having the lowest proportion of Uplands (12 %), then Wales (48 %), and Scotland having the largest overall area of both true and intermediate Uplands (73 %).

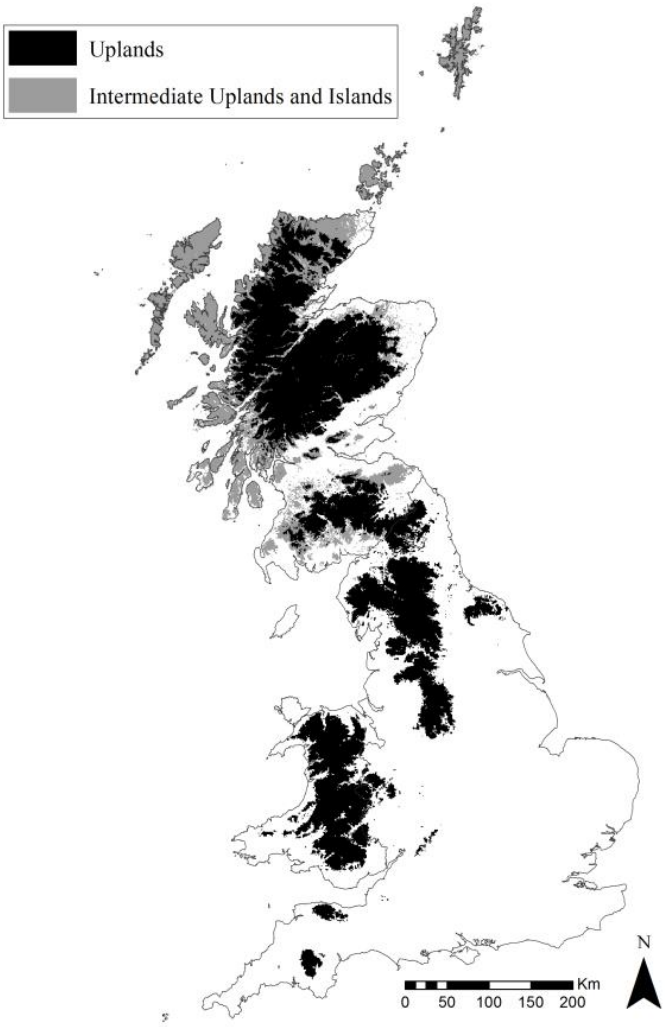
The total area of Uplands given by Bunce (1987) was 7.7 m ha (39 %), of which only 4.6 m ha (23 %) are upland vegetation. The total figure is consistent with the figure of 40 % given by Royal Society for the Protection of Birds (RSPB, 2007) and that of over 30 % quoted by Monbiot (2017a). There is, therefore, a broad consensus of the general area, but it is the detail and comprehensive coverage which is lacking - hence the present article.

Table 1: Percentage and area of upland landscapes in Great Britain, England, Scotland and Wales derived from the Environmental Land Classes of Bunce *et al.* (2007) with upland characteristics.

England: classes 16 and 20-23, Scotland true Uplands 29-32, Intermediate Uplands and Islands 27-28, 36-40 and Wales: classes 17-19 and 45.

	%		Area ('000a ha)	
Great Britain	**38 %**		8791	
England	12 %		1574	
Scotland (true uplands)	38 %	73 %	3203	6190
Scotland (intermediate uplands and islands)	35 %		2987	
Wales	48 %		1027	

Fig. 1: Distribution of upland Land Classes in Britain, derived from the ITE Land Classification (Bunce *et al.*, 2007)



Habitats

The figures given in Table 2 are comparable with those given by JNCC (2012), for Upland Heath: England 243 k ha (present paper 270 k ha), Scotland 778 k ha (826 k ha), Wales: 80 k ha (112 k ha). The differences are wide for bog because of different definitions, especially in regard to blanket bog, to the extent that comparisons are not useful.

Table 2 shows that the bog habitat is the most widespread in GB covering 2062 k ha although it is not so extensive in England and Wales because, as Table 1 shows, Scotland has the major proportion of upland landscapes and is dominated by bog. Acid grassland covers 1442 k ha and dwarf shrub heath 1208 k ha, which is the *Calluna vulgaris* (ling heather) dominated habitat often used by grouse (see Table 3). Although heather moorland is often associated in public perception with the Uplands, there is actually more acid grassland overall than dwarf shrub heath, reflecting the impact of grazing.

The importance of the influence of man is further highlighted by the proportion of highly managed habitats (grasslands and coniferous forest) which is almost 50 %, indicating that the Uplands are not the wilderness that they are often perceived to be (Smith *et al*., 2012) and also showing the importance of objective figures. The most comparable complex of habitats elsewhere in Europe is in Western Norway, as reflected in the Environmental classes of Metzger *et al.* (2005), although there is more semi-natural birch and pine woodland, and less grassland, than in Scotland.

Table 2 shows that, surprisingly, in England the typical upland habitats bog and dwarf shrub heath together occupy a smaller area than the two generally lowland habitats improved and neutral grassland. This pattern is even more pronounced in Wales where improved grassland is the most abundant Broad Habitat reflecting the less rugged terrain allowing vehicular access for fertilizer application. Both Scotland and Wales have over 10 % coniferous forest indicating the impact of plantations of exotic conifers as described by Bunce *et al*. (2014b). Another surprising figure is the relatively low cover of bracken in all three countries, which is often considered to be a major problem. This could be because it is often present as part of the acid grassland habitat, not in sufficient density (>95 %) to be recorded as a Broad Habitat in its own right. Arable land only features in England, whereas the relatively high figure for standing water in Wales and Scotland reflects the presence of reservoirs in the former and natural lochs in the latter.

Table 2: The top ten Broad Habitats of the UK Biodiversity Action Plan (Jackson, 2000) in upland landscapes in England, Scotland and Wales. Ranked in terms of extent in each country. Percentage in column 1, area in column 2.

Great Britain	%	Area ('000s ha)	England	%	Area ('000s ha)	Scotland	%	Area ('000s ha)	Wales	%	Area ('000s ha)
Bog	25 %	2061.6	Acid grassland	23 %	349.4	Bog	33 %	1888.1	Improved grassland	26 %	262.7
Acid grassland	17 %	1442.0	Dwarf shrub heath	18 %	269.7	Acid grassland	16 %	900.9	Acid grassland	19 %	191.7
Dwarf shrub heath	15 %	1207.7	Neutral grassland	16 %	237.1	Dwarf shrub heath	14 %	825.8	Dwarf shrub heath	11 %	112.3
Coniferous woodland	11 %	949.1	Improved grassland	15 %	224.9	Coniferous woodland	14 %	774.6	Neutral grassland	10 %	104.2
Improved grassland	10 %	814.7	Bog	9 %	133.9	Improved grassland	6 %	327.1	Coniferous woodland	10 %	102.4
Neutral grassland	6 %	525.0	Coniferous woodland	5 %	72.1	Neutral grassland	3%	183.8	Broadleaved, mixed and yew woodland	9 %	89.6
Broadleaved, mixed and yew woodland	3 %	250.1	Bracken	4 %	56.7	Fen, marsh and swamp	3 %	167.0	Bog	4 %	39.6
Fen, marsh and swamp	3 %	239.1	Fen, marsh and swamp	3 %	50.4	Broadleaved, mixed and yew woodland	2 %	120.2	Bracken	3 %	33.1
Bracken	2 %	198.4	Broadleaved, mixed and yew woodland	3 %	40.4	Bracken	2 %	108.6	Fen, marsh and swamp	2 %	21.7
Arable and horticultural	1 %	116.3	Arable and horticultural	2 %	34.1	Standing water and canals	1 %	81.9	Standing water and canals	1 %	14.8

Table 3: Percentage and area of grazing animals recorded in land parcels within the sample 1 km squares (Brown *et al.*, 2016).

Upland Zone	Animal	% upland area grazed	Area ('000ha)
Uplands (England)	Cattle	4.2	65.6
	Deer	0.1	1.1
	Grouse	7.1	112.1
	Sheep	39.9	628.4
	Any grazing animal	*45.3*	*713.1*
Intermediate uplands and Islands (Scotland)	Cattle	3.3	97.7
	Deer	11.8	352.5
	Grouse	2.3	69.5
	Sheep	42.5	1270.5
	Any grazing animal	*50.8*	*1516.7*
True uplands (Scotland)	Cattle	0.7	21.4
	Deer	37.4	1197.0
	Grouse	17.5	559.1
	Sheep	37.8	1209.9
	Any grazing animal	*56.5*	*1808.5*
Uplands (Wales)	Cattle	4.1	42.4
	Sheep	47.4	487.4
	Any grazing animal	*48.7*	*500.0*

The figures in Table 3 demonstrate the dominance of sheep grazing throughout the British Uplands and explain the dominance of acid grassland over dwarf shrub heath. The removal of sheep grazing and the subsequent shift from acid grassland to dwarf shrub heath species is discussed by Hill *et al*. (1992). The figures for cattle are low – being present in about 4 % of Uplands in England, Wales and the Scottish Intermediate Uplands, and under 1 % in the true Scottish Uplands. These figures emphasize the decline of hill cattle - even more so considering that valleys were included. Sheep were present in almost 50 % of the Uplands in all countries with Wales having the highest figure. Grouse are absent from Wales, very low in the intermediate Uplands in Scotland (2.3 %) but common in England (7.1 %) and highest in the Scottish Highlands (17.5 %), indicating the importance of shooting in this region. Red deer are absent from Wales, very low in England (0.1 %) but high in the intermediate Scottish Highlands (11.8 %) but highest in the true Highlands (37.4 %) which are the focus of deer stalking.

Table 4: Estimates of the length of linear features and percentages in the uplands of Great Britain, England, Scotland and Wales.

Feature	Great Britain		England		Scotland		Wales	
	Mean length (000s km)	*% of total linears in GB*	*Mean length (000s km)*	*% of total linears in England*	*Mean length (000s km)*	*% of total linears in Scotland*	*Mean length (000s km)*	*% of total linears in Wales*
Managed hedgerows	29.66	6.2 %	9.4	2.4 %	2.6	12.4 %	17.6	32.7 %
Walls	84.19	48.4 %	43.9	53.9 %	34.0	43.3 %	6.3	46.7 %
Lines of trees/relict hedges & fence	18.72	16.4 %	5.6	7.7 %	2.8	23.1 %	10.4	34.7 %
Lines of trees/relict hedges	14.9	13.1 %	5.6	6.8 %	4.0	30.2 %	5.4	28.5 %
Bank/grass strip	12.29	19.2 %	2.4	5.6 %	4.0	64.9 %	5.9	36.7 %
Fence	181.4	27.3 %	33.1	9.1 %	114.4	50.5 %	33.9	45.7 %

There is a surprising number of hedges as well as lines of trees/relict hedges in all three countries; these are important as landscape features as well as for biodiversity. Although these are not usually associated with the Uplands, they are present in the valleys linked to the two generally lowland habitats. Again, the comprehensive coverage of the present paper has identified the importance of these habitats in the Uplands. The lengths of wall (mainly drystone) are especially high in England (43.9 k km) and Scotland (34 k km) but there are also significant lengths in Wales (6.3 k km). Their extent is a unique feature of the British Uplands not found to the same extent elsewhere in Europe and represents an important resource for landscape character and biodiversity. Fences predominate in Wales (33.9 k km) and Scotland (114.4 k km) but are not so important for biodiversity.

Table 5: Frequency of the top ten classes of the Countryside Vegetation System (Bunce *et al.*, 1999a) in upland Great Britain. Numbers of plots that fell in that class and the overall percentage in upland Great Britain. (Note: this is all plot types)

Great Britain	No. plots *(total 5953)*	%
Rushy moorland grass/streamsides on peat soils	493	8 %
Moorland grass/heath on podzolic soils	306	5 %
Moorland grass	281	5 %
Cotton grass bog	245	4 %
Saturated bog	226	4 %
Moorland grass/bog on peaty gley/peat soils	223	4 %
Moorland grass/heath/bog	217	4 %
Rye-grass/bent grass grassland	212	4 %
Bracken/acid grassland	212	4 %
Fertile mixed grassland	186	3 %

Table 5 shows a rather different balance of vegetation in the landscape than the Broad Habitat extents, which are mapped as complete cover in each survey square (therefore including habitats such as open water and urban which do not have vegetation plots placed within them). It is important to note that the names of the Countryside Vegetation System (CVS) classes are only convenient labels to help interpretation of the 100 classes, which are determined on the basis of multivariate analysis (TWINSPAN, Hill & Šmilauer (2005)) of the complete species composition in each plot. They are, therefore, independent of the habitat classes. A full discussion of the results is given in Bunce *et al*. (1999a).

The most abundant class (rushy moorland grass/streamsides on peat soils) reflects the dominance of wetness in determining the composition of the vegetation and occurs by springs and seepages as well as beside streams. The next three classes (moorland grass/heath on podzolic soils, moorland grass, cotton grass bog) show the strong relationship of much of the vegetation in the Uplands with bogs in a broad context, reflecting the difficulty of defining bogs when mapping discrete polygons. The next two classes (saturated bog, moorland grass/bog on peaty gley/peat soils) show the complex intergrades of moorlands, heaths and bogs that are abundant in the Uplands, which are widespread. Two of the last three classes confirm the extent of vegetation with lowland affinities shown in the habitat results (rye-grass/bent grass grassland, fertile mixed grassland). The other class (bracken/acid grassland) indicates that there is an intergrade between acid grassland and bracken, with the abundance of bracken changing, depending on management.

Table 6: The top ten most frequent species in the dispersed random 200 m square (main) plots within the sample squares in England, Scotland and Wales (excluding bryophytes and lichens).

	Great Britain	No. of plots	Freq.	England	No. of plots	Freq.	Scotland	No. of plots	Freq.	Wales	No. of plots	Freq.
1	*Potentilla erecta*	666	59 %	*Galium saxatile*	152	60 %	*Potentilla erecta*	499	77 %	*Agrostis capillaris*	176	75 %
2	*Calluna vulgaris*	658	58 %	*Deschampsia flexuosa*	142	56 %	*Calluna vulgaris*	472	73 %	*Holcus lanatus*	128	54 %
3	*Molinia caerulea*	512	45 %	*Calluna vulgaris*	135	53 %	*Molinia caerulea*	387	60 %	*Cerastium fontanum*	117	50 %
4	*Anthoxanthum odoratum*	506	45 %	*Anthoxanthum odoratum*	135	53 %	*Trichophorum cespitosum*	333	51 %	*Trifolium repens*	116	49 %
5	*Agrostis capillaris*	502	44 %	*Agrostis capillaris*	133	53 %	*Erica tetralix*	328	51 %	*Anthoxanthum odoratum*	108	46 %
6	*Galium saxatile*	492	43 %	*Vaccinium myrtillus*	129	51 %	*Eriophorum angustifolium*	297	46 %	*Lolium perenne*	106	45 %
7	*Deschampsia flexuosa*	490	43 %	*Nardus stricta*	124	49 %	*Juncus squarrosus*	292	45 %	*Galium saxatile*	104	44 %
8	*Vaccinium myrtillus*	488	43 %	*Festuca ovina agg.*	123	49 %	*Narthecium ossifragum*	284	44 %	*Vaccinium myrtillus*	96	41 %
9	*Juncus squarrosus*	454	40 %	*Holcus lanatus*	112	44 %	*Anthoxanthum odoratum*	263	41 %	*Ranunculus repens*	90	38 %
10	*Nardus stricta*	445	39 %	*Juncus effusus*	104	41 %	*Vaccinium myrtillus*	263	41 %	*Deschampsia flexuosa*	86	37 %

This table (6) shows that *Potentilla erecta* (tormentil), a species of acid soils, is the most frequent species in upland Britain, with four other acid grassland plants in the top ten. Table 6 provides further detail of the composition of the vegetation and habitats and confirms the differences in the three countries as indicated by the composition of habitats. The species that have the highest frequency in Wales are all grassland plants, with the exception of *Vaccinium myrtillus* (bilberry), and also includes three species from fertile soils (*Trifolium repens* (white clover), *Lolium perenne* (perennial rye-grass), *Ranunculus repens* (creeping buttercup)). By contrast, the widespread species in England include none from fertile soils with the exception of *Holcus lanatus* (Yorkshire fog) and are otherwise evenly divided between those of acid grasslands and heathland. In Scotland, all the species are from acid soils and wet peats with three species of dwarf shrubs *(Calluna vulgaris, Erica tetralix, Vaccinium myrtillus)*, emphasizing the extreme oceanic nature of the region and its similarity with Norway.

In terms of species coverage, although not appearing in a high position in the frequency table, *Calluna vulgaris* (Ling heather) has the highest cover overall in plots in the British Uplands (mean per plot, 14.8 %). A series of histograms presented in the Supplementary Material (S1) show the variation in cover within plots of the species in Table 7. Cover of *Calluna* in particular is shown to dominate some plots. *Calluna* dominates in England (14.7 %) and Scotland (18.7 %), demonstrating the importance of this species overall (although it is only in tenth position in Wales). The extent of this species is not reflected in the habitat or vegetation tables but confirms the public perception of the purple colour of the British Uplands during flowering of this species in summer. *Lolium perenne* (perennial rye-grass) appears as the dominant species in Wales, in second position in England and fifth in Scotland, further confirming the importance of fertile fields in the Uplands. Otherwise, Scotland has the most bog species for example *Molinia caerulea* (purple moor grass) and *Tricophorum caespitosum* (deer grass). By contrast, England has the most acid grassland species, for example *Agrostis capillaris* (bent grass) and *Deschampsia flexuosa* (wavy hair grass), whereas Wales has the most bracken (*Pteridium aquilinum*).

Table 7: Average percentage cover of the top ten species (with the greatest overall cover) recorded in the dispersed random 200 m square (main) plots within the sample squares in Great Britain (excluding bryophytes and lichens)

	Great Britain	Mean cover per plot	England	Mean cover per plot	Scotland	Mean cover per plot	Wales	Mean cover per plot
1	*Calluna vulgaris*	14.8	*Calluna vulgaris*	14.7	*Calluna vulgaris*	18.7	*Lolium perenne*	17.7
2	*Molinia caerulea*	10.7	*Lolium perenne*	9.0	*Molinia caerulea*	13.5	*Agrostis capillaris*	12.6
3	*Lolium perenne*	8.1	*Vaccinium myrtillus*	8.3	*Trichophorum cespitosum*	9.5	*Molinia caerulea*	7.2
4	*Trichophorum cespitosum*	5.7	*Agrostis capillaris*	7.5	*Picea sitchensis*	5.9	*Pteridium aquilinum*	5.8
5	*Agrostis capillaris*	5.6	*Deschampsia flexuosa*	7.2	*Lolium perenne*	4.3	*Vaccinium myrtillus*	5.7
6	*Picea sitchensis*	5.0	*Molinia caerulea*	7.0	*Eriophorum vaginatum*	4.0	*Nardus stricta*	5.6
7	*Nardus stricta*	4.5	*Pteridium aquilinum*	5.9	*Nardus stricta*	3.7	*Picea sitchensis*	5.2
8	*Vaccinium myrtillus*	4.0	*Nardus stricta*	5.8	*Eriophorum angustifolium*	3.3	*Trifolium repens*	4.9
9	*Pteridium aquilinum*	3.8	*Holcus lanatus*	5.5	*Holcus lanatus*	2.6	*Holcus lanatus*	4.8
10	*Holcus lanatus*	3.7	*Festuca ovina agg.*	5.2	*Deschampsia flexuosa*	2.6	*Calluna vulgaris*	4.4

The role of grazing animals in the Uplands

Sheep are the most widespread domestic grazing animal throughout the British Uplands as shown in Table 3, and not only have had a major role in the formation of the present composition of the vegetation, but also in maintaining its current condition. The total breeding flock of sheep, as defined by the Annual Statistics for sheep and lambs on December 1 (Department for Environment Food & Rural Affairs, 2018), has declined somewhat since its peak of 31 million in 1998 to 22.0 million in 2008, recovering slightly to 23.3 million in 2017. Whilst these figures are for the whole country, not just the Uplands, they give an indication of the extent of the grazing pressure. (It would be a major exercise to extract separate figures for the Uplands, and Parish figures are often misleading (for example Scottish Government (2016)), as the sheep may all be in the valleys and may not use the hills at all, as in Ennerdale in the English Lake District). Fuller & Gough (1999) discuss the changes in sheep numbers in the Uplands and emphasize that the effect of grazing on the structure of semi-natural vegetation reduces their value for birds.

Table 3 shows that red deer are mainly in Scotland, although there are local herds in the Lake District and Exmoor. Whilst estimates of numbers have to be treated with caution, Clutton-Brock *et al*. (2004) indicate that the population was around 150,000 in 1960 but had risen to 400,000 by 2004. Flyn (2017) quotes a current figure of 1.5 million, which seems high, although it is generally accepted that numbers are still increasing. Flyn (2017) also suggests the reasons behind the increases, which are partly due to the absence of any predator but mainly because shooting female deer is not profitable. The call for a mass cull is very controversial and is unlikely to take place at present for a variety of reasons, varying from the views of landowners and professional hunters to the attitude of the public. What is beyond doubt is that currently the number of deer is having a negative influence on the vegetation and prohibit tree regeneration, as discussed by Bunce *et al*. (2014b). Some estates (for example, the National Trust for Scotland at Mar Lodge), are now reducing numbers and regeneration is resulting (Gill & Morgan, 2010). Other estates, such has Abernethy and Glenfeshie are also reducing numbers, a trend followed by publicly owned land (such as that of Forest Enterprise).

There is much discussion about the possibility that Brexit may result in the loss of sheep farming in the Uplands, leading to the suggestion that the hills will become dense scrub. However, this consequence may be variable. Hill *et al.* (1992), show that although change is rapid at first, eventually competitive species such as *Calluna vulgaris* and *Nardus stricta* take over the sward to the extent that other species are then unable to colonise. Also, there is an absence of tree seed in many upland landscapes (for example the Southern Uplands of Scotland). Table 7 shows that *Calluna*, *Vaccinium, Eriophorum, Molinia, Nardus* and *Tricophorum* are the major cover species in the Uplands, which can all form a dense impenetrable sward, resistant to change, thereby forming plagioclimaxes, as seen in forest rides.

The expansion of tree cover is therefore likely to be variable depending on soil type, altitude and the past history of the vegetation. For example, abundant tree colonisation by Scots pine (*Pinus sylvestris*) is shown in the pictures of Bain (2013) adjacent to native pinewoods, and birch (*Betula* species) in the picture of Ennerdale in Bunce *et al*. (2014b). There is also a history of rapid tree regeneration in Norway and the Pyrenees, although the soils are different from most British Uplands.

Further research is needed to establish whether sheep are still transferred to the higher mountain land, as there is some evidence from the English Lake District and the Cheviots that this practice is in decline.

Genetic diversity in the Uplands.

The Uplands contain important resources of genetic diversity in the wide range of native cattle and sheep breeds present. For example, cattle are represented by Welsh Black, Galloway and Scottish Highland breeds, all of which originated in the Uplands. There are also many sheep breeds often localised in their occurrence (such as Welsh Mountain (Wales), Herdwick (The Lake District), Scottish Blackface (Scottish mountains) and Soay sheep (the Hebrides)). Wild goats are also present, for example in Snowdonia and The Cheviots, and local breeds of ponies in Dartmoor and Snowdonia. Both the latter have been notified as threatened breeds (Murray, 2007).

DISCUSSION

There seems little doubt that major changes in upland habitats will occur in the future, as they have throughout history, but particularly following Brexit. The policy issues driving some of these changes are addressed in the second part of this paper (Bunce *et al.*, 2018). Changes are likely to be very different at regional levels, with isolation from markets and landscape structure being key factors. For example, the rounded hills of central Wales could be suitable for large scale ranching agriculture run by a few farms, whereas the rugged landscapes of North West Scotland could not be managed as large units. Elsewhere in Europe, for example in the Pyrenees (Baudry & Bunce, 1991), similar heterogeneous landscapes have largely been abandoned, a process that to date has not happened in Britain, although the first signs can be seen in the far North West of Scotland.

Throughout the upland region, social structure is likely to have a major influence, as is recognised in the definition of Less Favoured Areas, and in common with mountainous regions elsewhere in Europe. On the one hand, there is an aging population, with young people not being willing to take up the hard lifestyle, but there is also the effect of isolation of communities especially in the Highlands and Islands of Scotland. The changing social structure is well shown in the Shetlands with the widespread abandonment of subsistence crofting, accelerated by the financial influence of the oil industry (Wood & Bunce, 2016).

Affluent sections of society have also had a major influence on rural housing in the case of second home ownership, indirectly linked to recreation. Under current economic conditions, this process is likely to continue and could be beneficial to the environment, as such people often wish to maintain traditional landscapes. The social impact of rising house prices is, however, a different matter and is perceived to be leading to a decline in schools and other local services, and changes in social structure (discussed in Hodge & Monk (2004); Stockdale *et al*. (2000)).

Finally, climate change could have a range of influences from reducing sub-arctic vegetation on the one hand, to increasing the potential for crop growth in fields in the valleys on the other. There are also likely to be major regional differences because of the wide climatic gradient in GB.

ACKNOWLEDGMENTS

The paper is the culmination of many discussions with colleagues, and Mike Daniels made some helpful comments regarding the manuscript. Thanks also go to Geoffrey Griffiths and an anonymous reviewer for suggesting useful improvements, and to Gaynor Greenwood for help with references.

REFERENCES

Averis, A., Averis, B., Birks, J., Horsfield, D., Thompson, D., & Yeo, M. (2004). *Illustrated guide to British upland vegetation*. Exeter: Joint Nature Conservation Committee, Pelagic Publishing.

Bain, C. (2013). *The Ancient Pinewoods of Scotland* Dingwall: Sandstone Press Ltd.

Baudry, J., & Bunce, R. G. H. (Eds.). (1991). *Land abandonment and its role in conservation. (Options Mediterraneennes. Serie A. Seminaires Mediterraneennes no.15)*: Zaragoza: Centre International de Hautes Etudes Agronomiques Mediterraneennes.

Berry, P., Dawson, T., Harrison, P., & Pearson, R. (2002). Modelling potential impacts of climate change on the bioclimatic envelope of species in Britain and Ireland. *Global ecology and biogeography, 11*(6), 453-462. doi: https://doi.org/10.1111/j.1466-8238.2002.00304.x.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2014a). *Countryside Survey 2007 estimates of Broad Habitat areas in Great Britain*. doi:10.5285/f03cba75-8bca-4679-ae2a-77a9fcbd4df3: NERC Environmental Information Data Centre.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2014b). *Countryside Survey 2007 estimates of linear feature lengths in Great Britain*. doi:10.5285/e687330b-a0f7-45a1-b58c-398e67da3028: NERC Environmental Information Data Centre.

Brown, M. J., Bunce, R. G. H., Carey, P. D., Chandler, K., Crowe, A., Maskell, L. C., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., Wood, C. M., & Wright, S. M. (2016). *Landscape area data 2007* [Countryside Survey]. doi:10.5285/bf189c57-61eb-4339-a7b3-d2e81fdde28d: NERC Environmental Information Data Centre.

Bunce, R., Bogers, M., Ortega, M., Morton, D., Allard, A., Prinz, M., Peterseil, J., Elena-Rossello, R., & Jongman, R. (2012). *Conversion of European habitat data sources into common standards*: Alterra.

Bunce, R. G. H. (1978). *UK Ecological Survey*. Handbook of Field Methods. Grange-over-Sands: Insitute of Terrestrial Ecology.

Bunce, R. G. H. (1979). Ecological survey of Britain *Annual Report of the Institute of Terrestrial Ecology* (Vol. 1978, pp. 74-75). Cambridge: Institute of Terrestrial Ecology.

Bunce, R. G. H. (1987). *The extent and composition of upland areas in Great Britain*. Paper presented at the In: Agriculture and conservation in the hills and uplands, edited by M.Bell & R.G.H.Bunce, 19-21. (ITE symposium no.23). Grange-over-Sands: Institute of Terrestrial Ecology.

Bunce, R. G. H. (1999). A framework for the analysis of cultural landscapes in Europe *In: Nature and culture in landscape ecology: experience for the 3rd millenium, edited by P. Kovar, 69-73. CZ-IALE (Regional Organisation of the International Association for Landscape Ecology of the Czech Republic).*

Bunce, R. G. H., Barr, C. J., Clarke, R. T., Howard, D., & Scott, A. (2007). *ITE Land Classification of Great Britain 2007.* doi:10.5285/5f0605e4-aa2a-48ab-b47c-bf5510823e8f: NERC Environmental Information Data Centre.

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., & Howard, D. C. (1996). The ITE Land Classification: providing an environmental stratification of Great Britain. *Environmental Monitoring and Assessment, 39*, 39-46. doi: 10.1007/978-94-009-1653-1_5.

Bunce, R. G. H., Barr, C. J., Gillespie, M. K., Howard, D. C., Scott, W. A., Smart, S. M., Van de Poll, H. M., & Watkins, J. W. (1999a). *Vegetation of the British countryside - the Countryside Vegetation System*. Grange-over-Sands Institute of Terrestrial Ecology. (ECOFACT Volume 1).

Bunce, R. G. H., Carey, P. D., Maskell, L. C., Norton, L. R., Scott, R. J., Smart, S. M., & Wood, C. M. (2014a). *Countryside Survey 2007 vegetation plot data*. https://doi.org/10.5285/57f97915-8ff1-473b-8c77-2564cbd747bc NERC Environmental Information Data Centre.

Bunce, R. G. H., Smart, S. M., van de Poll, H. M., Watkins, J. W., & Scott, W. A. (1999b). *Measuring change in British vegetation. ECOFACT volume 2*. Huntingdon Institute of Terrestrial Ecology.

Bunce, R. G. H., Wood, C. M., & Smart, S. M. (2018). The Ecology of British Upland Landscapes. II. The influence of policy on the current character of the Uplands and the potential for change. *Journal of Landscape Ecology* 11 (3), DOI: 10.2478/jlecol-2018-0016.

Bunce, R. G. H., Wood, C. M., Smart, S. M., Oakley, R., Browning, G., Daniels, M. J., Ashmole, P., Cresswell, J., & Holl, K. (2014b). The landscape ecological impact of afforestation on the British uplands and some initiatives to restore native woodland cover. *Journal of Landscape Ecology, 7*(2), 5-24. doi: https://doi.org/10.2478/jlecol-2014-0013.

Carey, P. D., Wallis, S., Chamberlain, P. M., Cooper, A., Emmett, B. A., Maskell, L. C., McCann, T., Murphy, J., Norton, L. R., Reynolds, B., Scott, W. A., Simpson, I. C., Smart, S. M., & Ullyett, J. M. (2008). *Countryside Survey: UK Results from 2007*. Lancaster: NERC/Centre for Ecology & Hydrology.

Clother, L., & Finch, E. (2010). Farming in the English uplands *Defra Agricultural Change and Environment Observatory Research Report No 20*. London.

Clutton-Brock, T., Coulson, T., & Milner, J. (2004). Red deer stocks in the Highlands of Scotland. *Nature, 429*(6989), 261. doi: https://doi.org/10.1038/429261a.

Davies, G. M., Kettridge, N., Stoof, C. R., Gray, A., Ascoli, D., Fernandes, P. M., Marrs, R., Allen, K. A., Doerr, S. H., & Clay, G. D. (2016). The role of fire in UK peatland and moorland management: the need for informed, unbiased debate. *Phil. Trans. R. Soc. B, 371*(1696), 20150342. doi: https://doi.org/10.1098/rstb.2015.0342.

Department for Environment Food & Rural Affairs (2018). *Annual statistics on the number of livestock in England and the UK in June and December*. Retrieved October 17, 2018, from https://www.gov.uk/government/statistical-data-sets/structure-of-the-livestock-industry-in-england-at-december.

European Commission (1992). Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora. *Brussels, Belgium*.

European Commission. (2018). *Less favoured areas scheme*. Retrieved October 17, 2018, https://ec.europa.eu/agriculture/rural-development-previous/2007-2013/less-favoured-areas-scheme_en.

Evans, D., Morrison-Bell, C., McCracken, D., & Thompson, D. (2017). The Uplands: Onwards and Upwards. *BES Bulletin, 48*(4), 46-47.

Flyn, C. (2017). Winterkill *Granta* (Vol. 142). London Granta Publications.

Fuller, R., & Gough, S. (1999). Changes in sheep numbers in Britain: implications for bird populations. *Biological Conservation, 91*(1), 73-89. doi: https://doi.org/10.1016/S0006-3207(99)00039-7.

Gill, R., & Morgan, G. (2010). The effects of varying deer density on natural regeneration in woodlands in lowland Britain. *Forestry, 83*(1), 53-63. doi: https://doi.org/10.1093/forestry/cpp031.

Hill, M. O., Evans, D. F., & Bell, S. A. (1992). Long-Term Effects of Excluding Sheep from Hill Pastures in North Wales. *Journal of Ecology, 80*(1), 1-13. doi: 10.2307/2261058.

Hill, M. O., & Šmilauer, P. (2005). *TWINSPAN for Windows version 2.3*. Wallingford: Centre for Ecology and Hydrology, University of South Bohemia.

Hodge, I., & Monk, S. (2004). The economic diversity of rural England: stylised fallacies and uncertain evidence. *Journal of Rural Studies, 20*(3), 263-272. doi: https://doi.org/10.1016/j.jrurstud.2003.11.004.

House, J. I., Orr, H. G., Clark, J. M., Gallego-Sala, A. V., Freeman, C., Prentice, I. C., & Smith, P. (2010). Climate change and the British Uplands: evidence for decision-making. *Climate Research, 45*, 3-12. doi: https://doi.org/10.3354/cr00982.

Jackson, D. (2000). *Guidance on the Interpretation of the Biodiversity Broad Habitat Classification* (Terrestrial and Freshwater Types): Definitions and the Relationship with Other Habitat Classifications (JNCC Report, No 307), Vol. Report no. 307, pp. 73pp. Published online, http://jncc.defra.gov.uk/page-2433: JNCC.

JNCC (2012). *Biodiversity Action Reporting System*. Retrieved October 17, 2018, from http://webarchive.nationalarchives.gov.uk/20120310132550/https://ukbars.defra.gov.uk//plans/national.asp?S=&L=1&O=&SAP=&HAP=&submitted=1.

Maddock, A. (2008). UK Biodiversity Action Plan; Priority Habitat Descriptions *UK Biodiversity Action Plan, 94pp*. Published online, http://jncc.defra.gov.uk/PDF/UKBAP_PriorityHabitatDesc-Rev2011.pdf.

Maskell, L. C., Norton, L. R., Smart, S. M., Carey, P. D., Murphy, J., Chamberlain, P. M., Wood, C. M., Bunce, R. G. H., & Barr, C. J. (2008). *Countryside Survey. Field Mapping Handbook*, CS Technical Report No. 1/07. Lancaster: Centre for Ecology & Hydrology.

Metzger, M. J., Bunce, R. G. H., Jongman, R. H. G., Mücher, C. A., & Watkins, J. W. (2005). A climatic stratification of the environment of Europe. *Global Ecology and Biogeography, 14*(6), 549-563. doi: 10.1111/j.1466-822X.2005.00190.x.

Monbiot, G. (2017a, 11th January 2017). Explanation of the Figures in Grim Reaping Retrieved October 17, 2018, from https://www.monbiot.com/2017/01/11/explanation-of-the-figures-in-grim-reaping/.

Monbiot, G. (2017b, 06/01/2017). The new year brings challenges – but here's to a hopeful 2017, *The Guardian*.

Murray, D. A. (2007). The pony's tale: native breeds under threat, *Independent*. Retrieved August 2, 2007, from https://www.independent.co.uk/environment/nature/the-ponys-tale-native-breeds-under-threat-5334361.html.

Norton, L. R., Maskell, L. C., Smart, S. M., Dunbar, M. J., Emmett, B. A., Carey, P. D., Williams, P., Crowe, A., Chandler, K., Scott, W. A., & Wood, C. M. (2012). Measuring stock and change in the GB countryside for policy: key findings and developments from the Countryside Survey 2007 field survey. *Journal of Environmental Management, 113*, 117-127. doi: 10.1016/j.jenvman.2012.07.030.

Pearsall, W. H. (1950). *Mountains and moorlands*. London: Collins.

Ratcliffe, D., & Thompson, D. (1988). The British uplands: their ecological character and international significance. *Ecological change in the uplands*, 9-36.

Reed, M., Bonn, A., Slee, W., Beharry-Borg, N., Birch, J., Brown, I., Burt, T., Chapman, D., Chapman, P., & Clay, G. (2009). The future of the uplands. *Land use policy, 26*, S204-S216. doi: https://doi.org/10.1016/j.landusepol.2009.09.013.

RSPB (2007). *The Uplands*. Time to Change? Sandy, Bedfordshire.

Scottish Government (2016). Scottish Agricultural Census. Retrieved January 7, 2016, from http://www.gov.scot/Topics/Statistics/Browse/Agriculture-Fisheries/PubScottishCensus.

Sheail, J., & Bunce, R. G. H. (2003). The development and scientific principles of an environmental classification for strategic ecological survey in the United Kingdom. *Environmental Conservation, 30*(2), 147-159. doi: 10.1017/S0376892903000134.

Smith, E. W., Crockett, J., & McCoard, S. (2012). Public perception survey of wildness in Scotland: Report for Loch Lomond & The Trossachs National Park Authority *Cairngorms National Park Authority & Scottish Natural Heritage In Association With Research Now* Edinburgh: Scottish Natural Heritage.

Stockdale, A., Findlay, A., & Short, D. (2000). The repopulation of rural Scotland: opportunity and threat. *Journal of Rural Studies, 16*(2), 243-257. doi: https://doi.org/ 10.1016/S0743-0167(99)00045-5.

Wood, C. M., & Bunce, R. G. H. (2016). Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring. *Earth Syst. Sci. Data, 8*(1), 89-103. doi: 10.5194/essd-8-89-2016.

Wood, C. M., Bunce, R. G. H., Norton, L. R., Maskell, L. C., Smart, S. M., Scott, W. A., Henrys, P. A., Howard, D. C., Wright, S. M., Brown, M. J., Scott, R. J., Stuart, R. C., & Watkins, J. W. (2018). Ecological landscape elements: long-term monitoring in Great Britain, the Countryside Survey 1978–2007 and beyond. *Earth Syst. Sci. Data, 10*(2), 745-763. doi: 10.5194/essd-10-745-2018.

Wood, C. M., Smart, S. M., Bunce, R. G. H., Norton, L. R., Maskell, L. C., Howard, D. C., Scott, W. A., & Henrys, P. A. (2017). Long-term vegetation monitoring in Great Britain – the Countryside Survey 1978–2007 and beyond. *Earth Syst. Sci. Data, 9*(2), 445-459. doi: 10.5194/essd-9-445-2017.

Appendix ii: Technical work

This appendix contains descriptions of examples of technical work additional to those described in Section 3, in reference to Papers 2-7.

Woodlands (Paper 1)

Available resources (specific to the Paper 1 data):

- Field handbook
- MS Access database containing partial database
- Original field sheets

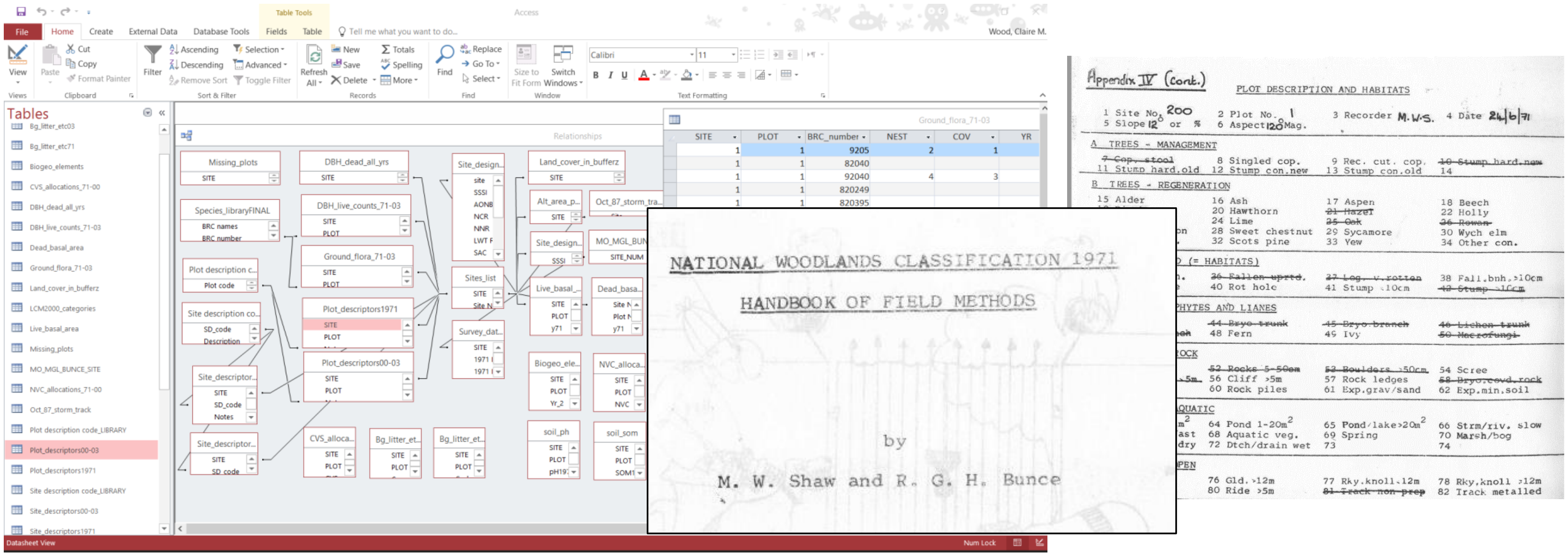
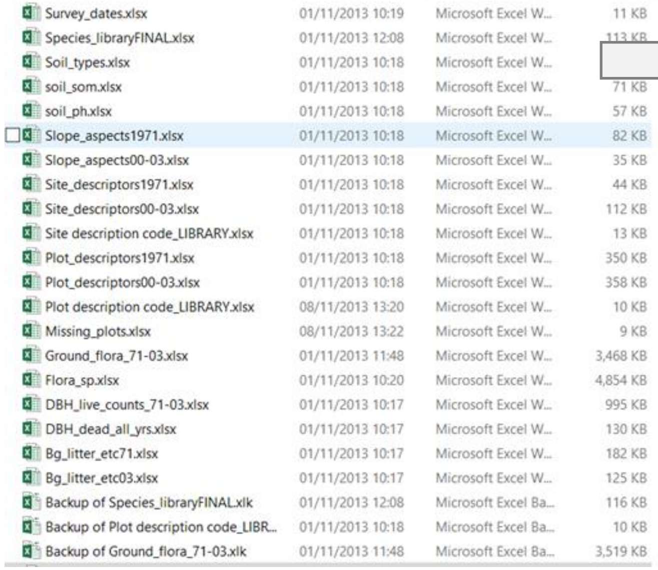


Figure ii. 1 Paper 1 resources

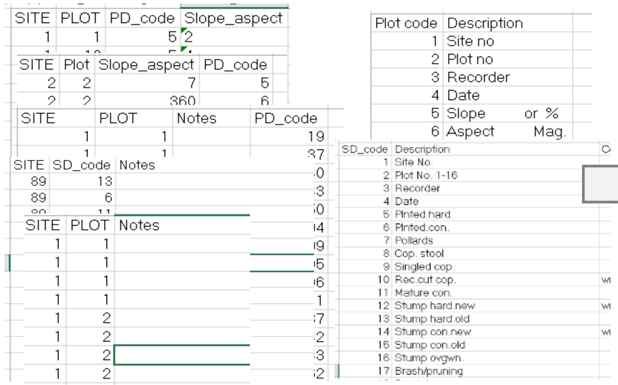
Table ii.1. Data tasks (specific to the Paper 1 data):

Quality/data issue	Example	Correction/Solution	Script ref.
Data from this survey was already available in a MS Access database/set of spreadsheets. However, the data did not match the schema design for other data sets (and various other issues, as described below, also existed).	Extant database, assembled following a re-survey in 2003.	Export tables to different schema with fewer tables for more efficient querying.	GBW1-8

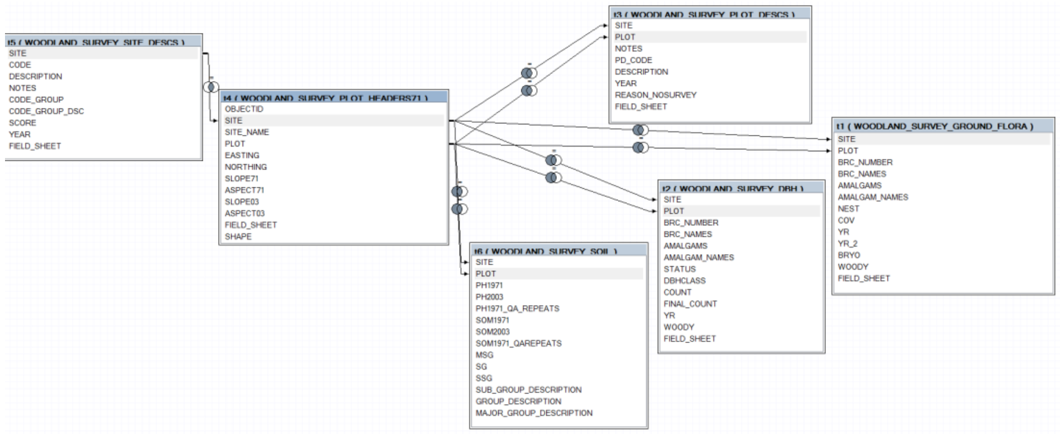


Name	Date modified	Type	Size
Survey_dates.xlsx	01/11/2013 10:19	Microsoft Excel W...	11 KB
Species_libraryFINAL.xlsx	01/11/2013 12:08	Microsoft Excel W...	113 KB
Soil_types.xlsx	01/11/2013 10:18	Microsoft Excel W...	
soil_som.xlsx	01/11/2013 10:18	Microsoft Excel W...	71 KB
soil_ph.xlsx	01/11/2013 10:18	Microsoft Excel W...	57 KB
Slope_aspects1971.xlsx	01/11/2013 10:18	Microsoft Excel W...	82 KB
Slope_aspects00-03.xlsx	01/11/2013 10:18	Microsoft Excel W...	35 KB
Site_descriptors1971.xlsx	01/11/2013 10:18	Microsoft Excel W...	44 KB
Site_descriptors00-03.xlsx	01/11/2013 10:18	Microsoft Excel W...	112 KB
Site description code_LIBRARY.xlsx	01/11/2013 10:18	Microsoft Excel W...	13 KB
Plot_descriptors1971.xlsx	01/11/2013 10:18	Microsoft Excel W...	350 KB
Plot_descriptors00-03.xlsx	01/11/2013 10:18	Microsoft Excel W...	358 KB
Plot description code_LIBRARY.xlsx	08/11/2013 13:20	Microsoft Excel W...	10 KB
Missing_plots.xlsx	08/11/2013 13:22	Microsoft Excel W...	9 KB
Ground_flora_71-03.xlsx	01/11/2013 11:48	Microsoft Excel W...	3,468 KB
Flora_sp.xlsx	01/11/2013 10:20	Microsoft Excel W...	4,854 KB
DBH_live_counts_71-03.xlsx	01/11/2013 10:17	Microsoft Excel W...	995 KB
DBH_dead_all_yrs.xlsx	01/11/2013 10:17	Microsoft Excel W...	130 KB
Bg_litter_etc71.xlsx	01/11/2013 10:17	Microsoft Excel W...	182 KB
Bg_litter_etc03.xlsx	01/11/2013 10:17	Microsoft Excel W...	125 KB
Backup of Species_libraryFINAL.xlk	01/11/2013 12:08	Microsoft Excel Ba...	116 KB
Backup of Plot description code_LIBR...	01/11/2013 10:18	Microsoft Excel Ba...	10 KB
Backup of Ground_flora_71-03.xlk	01/11/2013 11:48	Microsoft Excel Ba...	3,519 KB

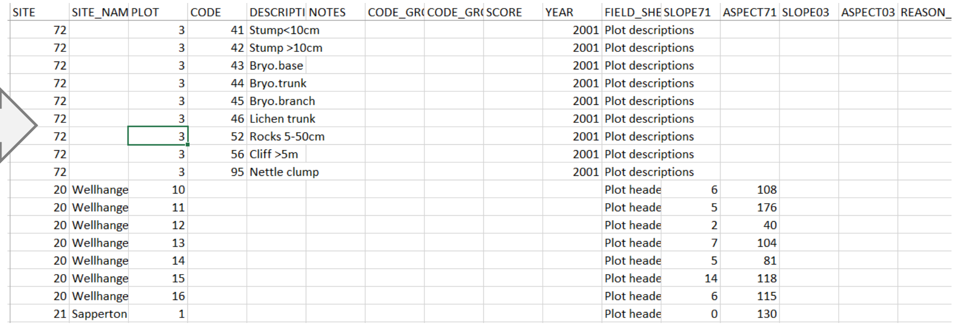
Example of existing set of tables




Woodlands_Survey_Flora_Data_1971_2001.csv
Woodlands_Survey_Site_Information_1971_2001.csv
Woodlands_Survey_Soil_Data_1971_2001.csv
Woodlands_Survey_Tree_Diameter_Data_1971_2001.csv



Final set of published tables and schema (showing tables as stored in the UKCEH Oracle database)



SITE	SITE_NAM	PLOT	CODE	DESCRIPTI	NOTES	CODE_GR(	CODE_GR(	SCORE	YEAR	FIELD_SHE	SLOPE71	ASPECT71	SLOPE03	ASPECT03	REASON_
72		3	41	Stump<10cm					2001	Plot descriptions					
72		3	42	Stump >10cm					2001	Plot descriptions					
72		3	43	Bryo.base					2001	Plot descriptions					
72		3	44	Bryo.trunk					2001	Plot descriptions					
72		3	45	Bryo.branch					2001	Plot descriptions					
72		3	46	Lichen trunk					2001	Plot descriptions					
72		3	52	Rocks 5-50cm					2001	Plot descriptions					
72		3	56	Cliff >5m					2001	Plot descriptions					
72		3	95	Nettle clump					2001	Plot descriptions					
20	Wellhange	10								Plot heade	6	108			
20	Wellhange	11								Plot heade	5	176			
20	Wellhange	12								Plot heade	2	40			
20	Wellhange	13								Plot heade	7	104			
20	Wellhange	14								Plot heade	5	81			
20	Wellhange	15								Plot heade	14	118			
20	Wellhange	16								Plot heade	6	115			
21	Sapperton	1								Plot heade	0	130			

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script ref.</th></tr>
<tr><td>Incorporate soils data</td><td>Avery soil codes existed in the database with no descriptions – description needed appending.
<table>
<tr><th>SITE</th><th>PLOT</th><th>MSG</th><th>SG</th><th>SSG</th></tr>
<tr><td>19</td><td>15</td><td>5</td><td>5.7</td><td>5.71</td></tr>
<tr><td>19</td><td>16</td><td>5</td><td>5.7</td><td>5.71</td></tr>
<tr><td>20</td><td>1</td><td>3</td><td>3.1</td><td>3.13</td></tr>
<tr><td>20</td><td>2</td><td>3</td><td>3.4</td><td>3.42</td></tr>
<tr><td>20</td><td>3</td><td>3</td><td>3.4</td><td>3.42</td></tr>
<tr><td>20</td><td>4</td><td>5</td><td>5.1</td><td>5.11</td></tr>
<tr><td>20</td><td>5</td><td>5</td><td>5.4</td><td>5.41</td></tr>
<tr><td>20</td><td>6</td><td>5</td><td>5.1</td><td>5.11</td></tr>
<tr><td>20</td><td>7</td><td>3</td><td>3.4</td><td>3.41</td></tr>
<tr><td>20</td><td>8</td><td>3</td><td>3.4</td><td>3.41</td></tr>
<tr><td>20</td><td>9</td><td>3</td><td>3.4</td><td>3.41</td></tr>
<tr><td>20</td><td>10</td><td>5</td><td>5.4</td><td>5.41</td></tr>
<tr><td>20</td><td>11</td><td>3</td><td>3.4</td><td>3.42</td></tr>
<tr><td>20</td><td>12</td><td>5</td><td>5.4</td><td>5.41</td></tr>
<tr><td>20</td><td>13</td><td>5</td><td>5.4</td><td>5.41</td></tr>
<tr><td>20</td><td>14</td><td>3</td><td>3.1</td><td>3.13</td></tr>
<tr><td>20</td><td>15</td><td>3</td><td>3.1</td><td>3.11</td></tr>
<tr><td>20</td><td>16</td><td>3</td><td>3.4</td><td>3.41</td></tr>
<tr><td>21</td><td>1</td><td>7</td><td>7.12</td><td>7.12</td></tr>
<tr><td>21</td><td>2</td><td>7</td><td>7.12</td><td>7.14</td></tr>
<tr><td>21</td><td>3</td><td>7</td><td>7.12</td><td>7.12</td></tr>
<tr><td>21</td><td>4</td><td>7</td><td>7.12</td><td>7.15</td></tr>
<tr><td>21</td><td>5</td><td>7</td><td>7.12</td><td>7.12</td></tr>
<tr><td>21</td><td>6</td><td>4</td><td>4.3</td><td>4.31</td></tr>
</table></td><td>Append soil type information for each plot as follows:
<table>
<tr><th>SITE</th><th>PLOT</th><th>MSG</th><th>SG</th><th>SSG</th><th>SUB_GROU...</th><th>GROUP_DE...</th><th>MAJOR_GR...</th></tr>
<tr><td>19</td><td>15</td><td>5</td><td>5.7</td><td>5.71</td><td>Typical argillic...</td><td>Argillic brown e...</td><td>Brown soils</td></tr>
<tr><td>19</td><td>16</td><td>5</td><td>5.7</td><td>5.71</td><td>Typical argillic...</td><td>Argillic brown e...</td><td>Brown soils</td></tr>
<tr><td>20</td><td>1</td><td>3</td><td>3.1</td><td>3.13</td><td>Brown rankers</td><td></td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>2</td><td>3</td><td>3.4</td><td>3.42</td><td>Grey rendzinas</td><td>Rendzinas</td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>3</td><td>3</td><td>3.4</td><td>3.42</td><td>Grey rendzinas</td><td>Rendzinas</td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>4</td><td>5</td><td>5.1</td><td>5.11</td><td>Typical brown c...</td><td>Brown calcareo...</td><td>Brown soils</td></tr>
<tr><td>20</td><td>5</td><td>5</td><td>5.4</td><td>5.41</td><td>Typical brown...</td><td>Brown earths (...</td><td>Brown soils</td></tr>
<tr><td>20</td><td>6</td><td>5</td><td>5.1</td><td>5.11</td><td>Typical brown c...</td><td>Brown calcareo...</td><td>Brown soils</td></tr>
<tr><td>20</td><td>7</td><td>3</td><td>3.4</td><td>3.41</td><td>Humic rendzinas</td><td>Rendzinas</td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>8</td><td>3</td><td>3.4</td><td>3.41</td><td>Humic rendzinas</td><td>Rendzinas</td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>9</td><td>3</td><td>3.4</td><td>3.41</td><td>Humic rendzinas</td><td>Rendzinas</td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>10</td><td>5</td><td>5.4</td><td>5.41</td><td>Typical brown...</td><td>Brown earths (...</td><td>Brown soils</td></tr>
<tr><td>20</td><td>11</td><td>3</td><td>3.4</td><td>3.42</td><td>Grey rendzinas</td><td>Rendzinas</td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>12</td><td>5</td><td>5.4</td><td>5.41</td><td>Typical brown...</td><td>Brown earths (...</td><td>Brown soils</td></tr>
<tr><td>20</td><td>13</td><td>5</td><td>5.4</td><td>5.41</td><td>Typical brown...</td><td>Brown earths (...</td><td>Brown soils</td></tr>
<tr><td>20</td><td>14</td><td>3</td><td>3.1</td><td>3.13</td><td>Brown rankers</td><td></td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>15</td><td>3</td><td>3.1</td><td>3.11</td><td>Humic rankers</td><td></td><td>Lithomorphic s...</td></tr>
<tr><td>20</td><td>16</td><td>3</td><td>3.4</td><td>3.41</td><td>Humic rendzinas</td><td>Rendzinas</td><td>Lithomorphic s...</td></tr>
<tr><td>21</td><td>1</td><td>7</td><td>7.12</td><td>7.12</td><td>Pelo-stagnogle...</td><td>Stagnogley soils</td><td>Surface-water...</td></tr>
<tr><td>21</td><td>2</td><td>7</td><td>7.12</td><td>7.14</td><td>Paleo-argillic st...</td><td>Stagnogley soils</td><td>Surface-water...</td></tr>
<tr><td>21</td><td>3</td><td>7</td><td>7.12</td><td>7.12</td><td>Pelo-stagnogle...</td><td>Stagnogley soils</td><td>Surface-water...</td></tr>
<tr><td>21</td><td>4</td><td>7</td><td>7.12</td><td>7.15</td><td>Sandy stagnogl...</td><td>Stagnogley soils</td><td>Surface-water...</td></tr>
<tr><td>21</td><td>5</td><td>7</td><td>7.12</td><td>7.12</td><td>Pelo-stagnogle...</td><td>Stagnogley soils</td><td>Surface-water...</td></tr>
<tr><td>21</td><td>6</td><td>4</td><td>4.3</td><td>4.31</td><td>Typical argillic...</td><td>Argillic pelosols</td><td>Pelosols</td></tr>
</table></td><td>GBW5</td></tr>
<tr><td>Quality checks to check for anomalous values</td><td>Examples of checks undertaken

Range checks:
pH must be between 1 and 14
Aspect must be between 0 and 360

Omission checks:
All coded data should have a corresponding description
All sites/plots are complete

Consistency checks:
The same values should be recorded the same in the data (i.e. ‘Beech’ not ‘BEECH’/’Fagus’)</td><td></td><td></td></tr>
</table>

Shetland (Paper 3)

Available resources:

- *Field handbook*
- *Contract Reports from 1974*
- *Field data sheets*
- *Laboratory notebooks*
- *Field maps*

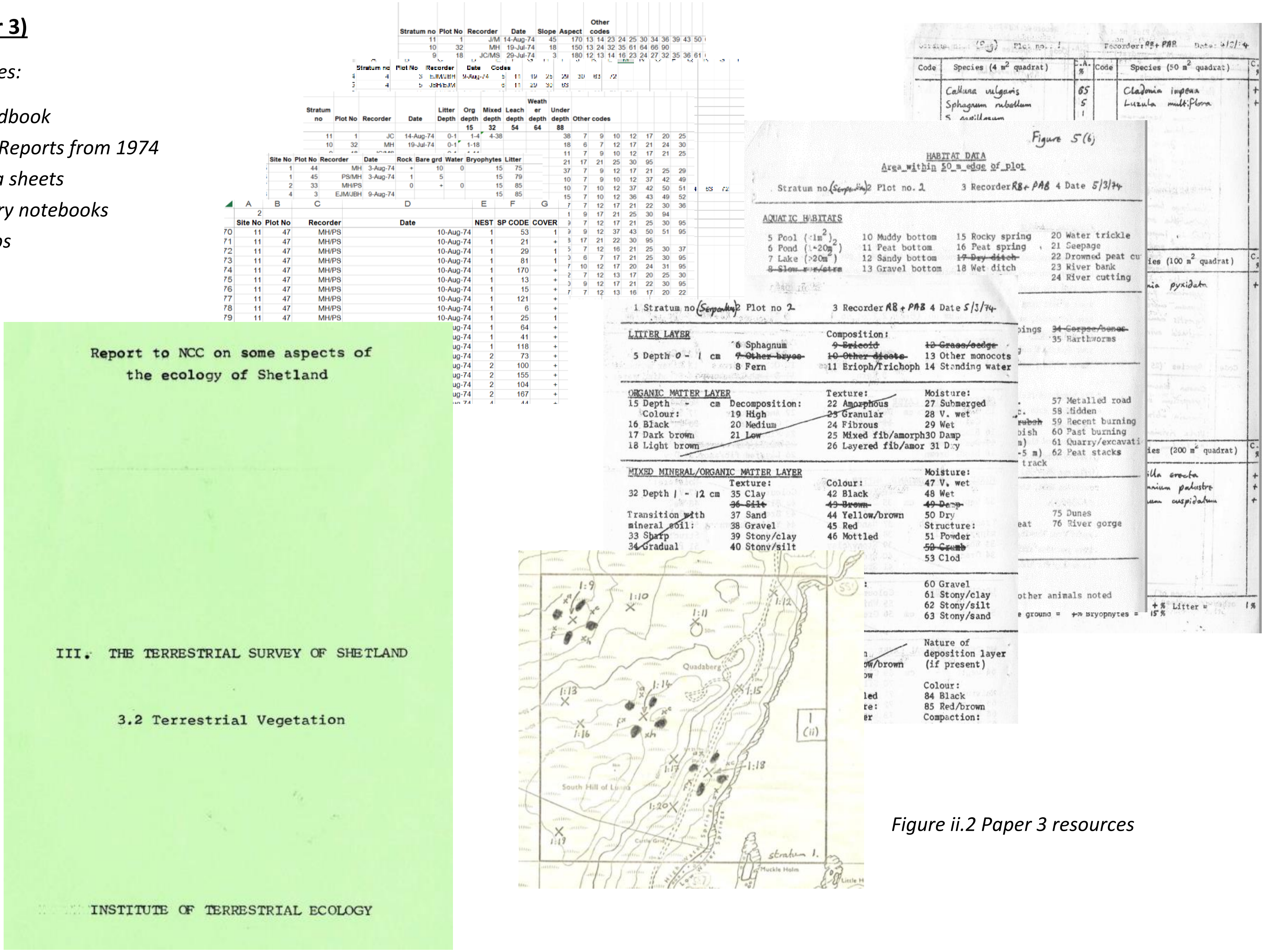
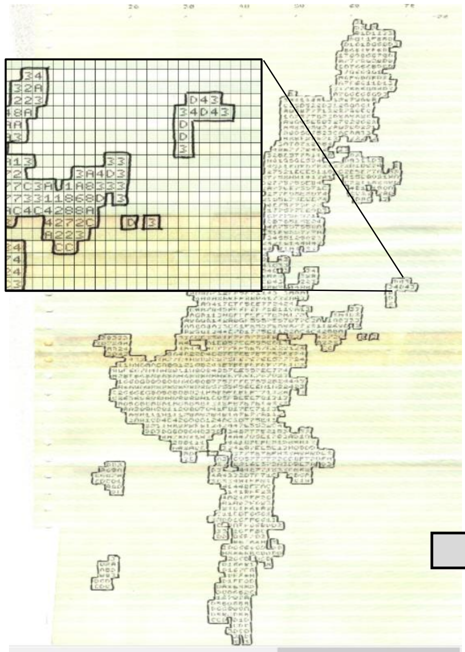
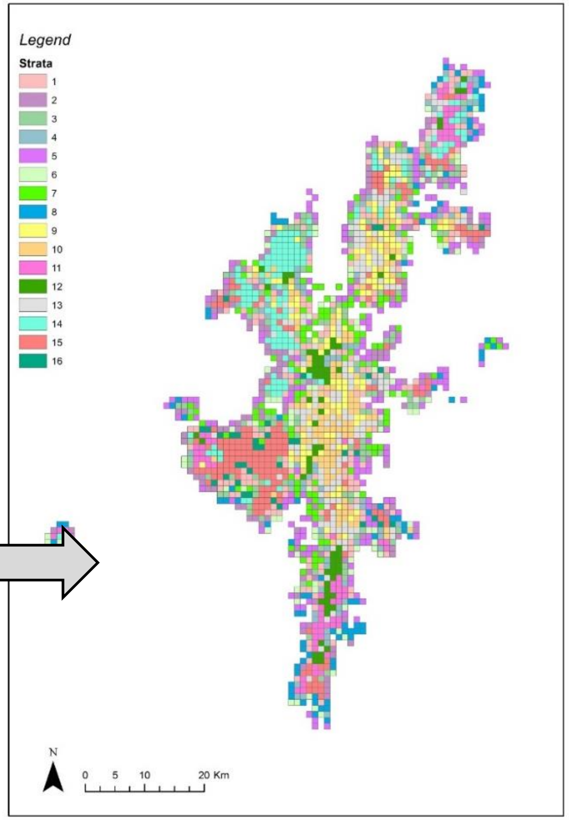


Figure ii.2 Paper 3 resources

Table ii.2. Data tasks (specific to the Paper 3 data):

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script ref.</th></tr>
<tr><td>Strata file missing from digital data set</td><td>Map still existing on paper:
</td><td>Locate strata map on paper. Digitise strata according to gridded layout. Locate strata information regarding original construction and descriptions in old contract report.

</td><td>N/A</td></tr>
</table>

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script ref.</th></tr>
<tr><td></td><td></td><td>
<table>
<tr><th>OS_TILE</th><th>SURVEY_STRATA</th><th>SURVEY_STRATA_SUB</th><th>MAP_STRATA_CODE</th><th>STRATA</th></tr>
<tr><td>HU3048</td><td>0</td><td></td><td>G</td><td>14</td></tr>
<tr><td>HU2147</td><td>0</td><td></td><td>A</td><td>2</td></tr>
<tr><td>HU3447</td><td>0</td><td></td><td>2</td><td>3</td></tr>
<tr><td>HU2946</td><td>0</td><td></td><td>8</td><td>15</td></tr>
<tr><td>HU4346</td><td>0</td><td></td><td>2</td><td>3</td></tr>
<tr><td>HU2845</td><td>0</td><td></td><td>8</td><td>15</td></tr>
<tr><td>HU3045</td><td>15</td><td>v</td><td>8</td><td>15</td></tr>
<tr><td>HU3345</td><td>0</td><td></td><td>3</td><td>5</td></tr>
<tr><td>HU5045</td><td>0</td><td></td><td>3</td><td>5</td></tr>
<tr><td>HU2644</td><td>0</td><td></td><td>A</td><td>2</td></tr>
<tr><td>HU3044</td><td>0</td><td></td><td>8</td><td>15</td></tr>
<tr><td>HU3344</td><td>0</td><td></td><td>3</td><td>5</td></tr>
<tr><td>HU4144</td><td>0</td><td></td><td>8</td><td>15</td></tr>
<tr><td>HU4444</td><td>0</td><td></td><td>1</td><td>1</td></tr>
<tr><td>HU4944</td><td>0</td><td></td><td>D</td><td>8</td></tr>
<tr><td>HU3043</td><td>0</td><td></td><td>1</td><td>1</td></tr>
<tr><td>HU3243</td><td>0</td><td></td><td>1</td><td>1</td></tr>
<tr><td>HU3943</td><td>1</td><td>v</td><td>1</td><td>1</td></tr>
<tr><td>HU4143</td><td>0</td><td></td><td>7</td><td>13</td></tr>
<tr><td>HU5243</td><td>5</td><td>iii</td><td>3</td><td>5</td></tr>
<tr><td>HU3242</td><td>0</td><td></td><td>C</td><td>6</td></tr>
<tr><td>HU3842</td><td>0</td><td></td><td>4</td><td>7</td></tr>
</table>
</td><td></td></tr>
<tr><td>Soil pH data missing</td><td>Data found in a soils lab book in the archive</td><td>Locate laboratory notebook in archive. Digitise data and add to data set.
<table>
<tr><th>SURVEY_STRATA</th><th>SURVEY_STRATA_SUB</th><th>PLOT</th><th>SOIL_PH</th></tr>
<tr><td>1</td><td>iv</td><td>36</td><td>4.5</td></tr>
<tr><td>1</td><td>iv</td><td>37</td><td>4.6</td></tr>
<tr><td>1</td><td>iv</td><td>38</td><td>5</td></tr>
<tr><td>1</td><td>iv</td><td>39</td><td>4.6</td></tr>
<tr><td>1</td><td>v</td><td>40</td><td>4.5</td></tr>
<tr><td>1</td><td>v</td><td>41</td><td>4.5</td></tr>
<tr><td>1</td><td>v</td><td>42</td><td>4.4</td></tr>
<tr><td>1</td><td>v</td><td>43</td><td>5.1</td></tr>
<tr><td>1</td><td>v</td><td>44</td><td>4.3</td></tr>
<tr><td>1</td><td>v</td><td>45</td><td>4.3</td></tr>
<tr><td>1</td><td>v</td><td>46</td><td>6</td></tr>
<tr><td>1</td><td>i</td><td>47</td><td>4.4</td></tr>
<tr><td>2</td><td>i</td><td>1</td><td>5.6</td></tr>
<tr><td>2</td><td>i</td><td>2</td><td>4.6</td></tr>
<tr><td>2</td><td>i</td><td>3</td><td>5.9</td></tr>
</table>
</td><td>N/A</td></tr>
<tr><td>Data not structured efficiently</td><td></td><td>Additional tasks closely followed the Procedures for Paper 2.</td><td>SPW1-3</td></tr>
</table>

‘Key Habitats’ (Paper 6)

Available resources:

- *Field sheets*
- *Photographs*
- *Field handbooks*
- *Contract reports*
- *Some data in old Oracle tables*

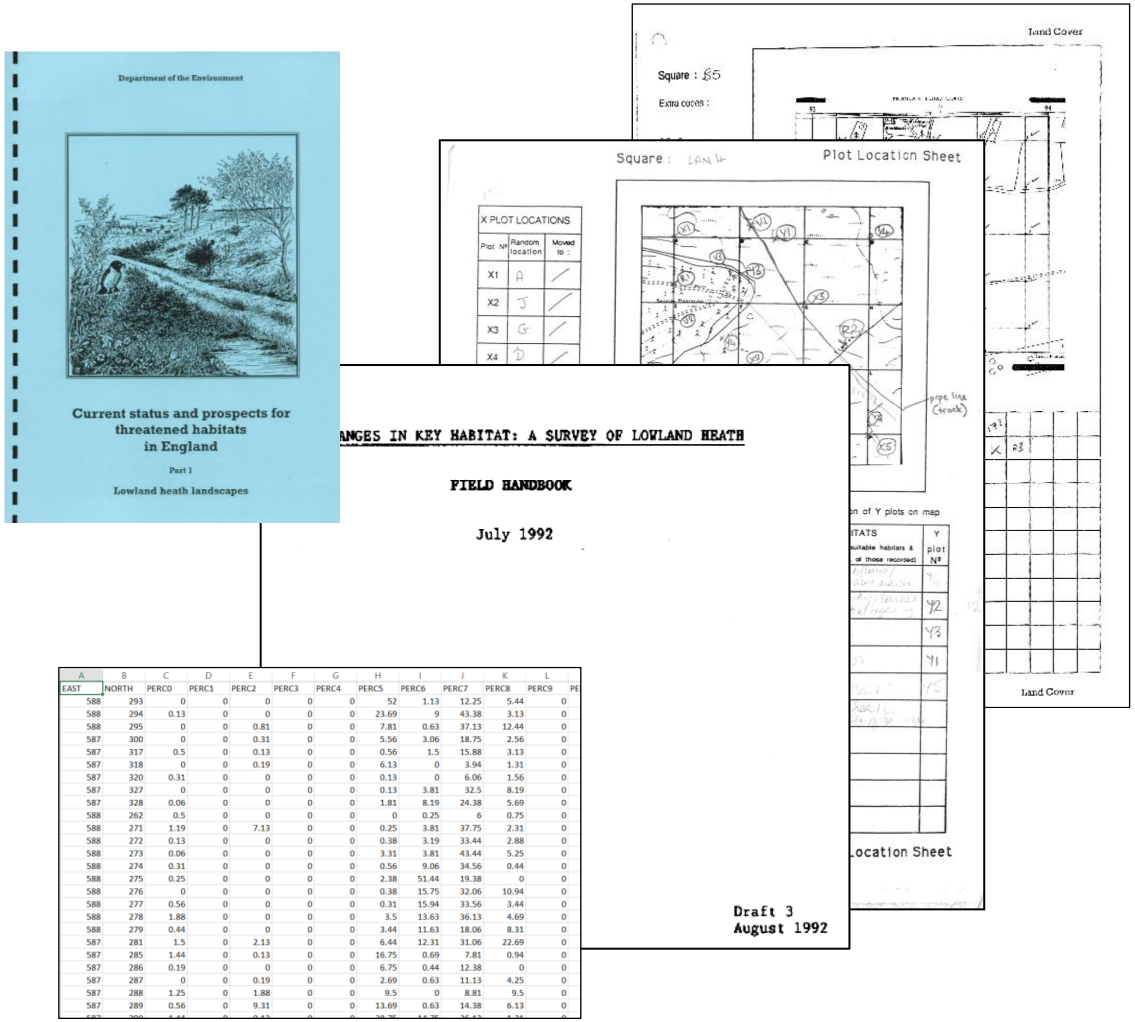
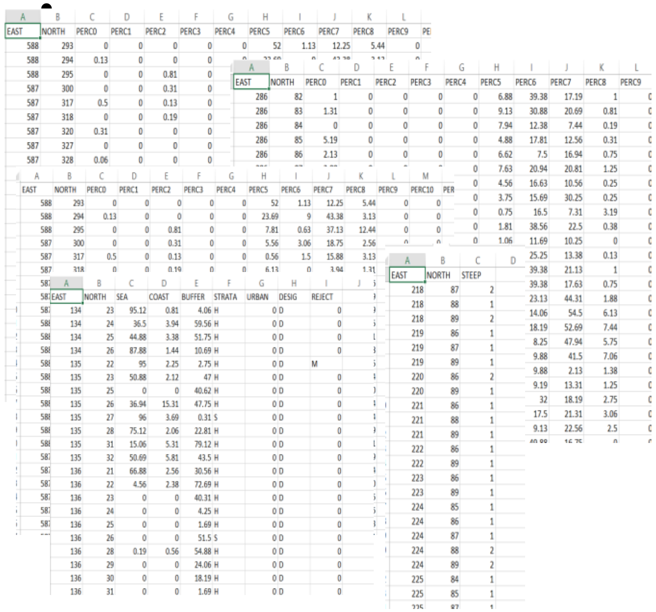
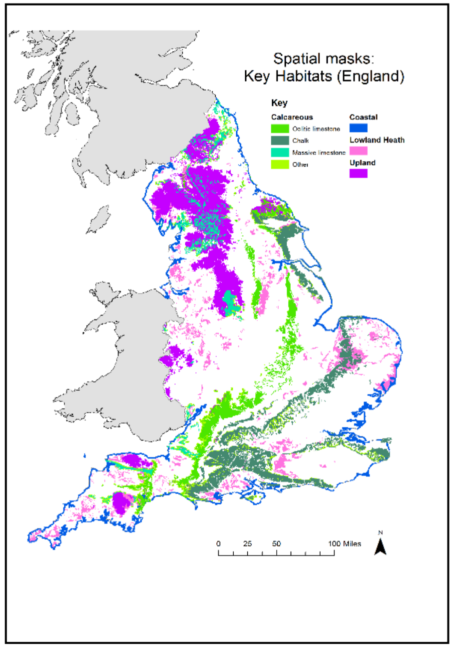
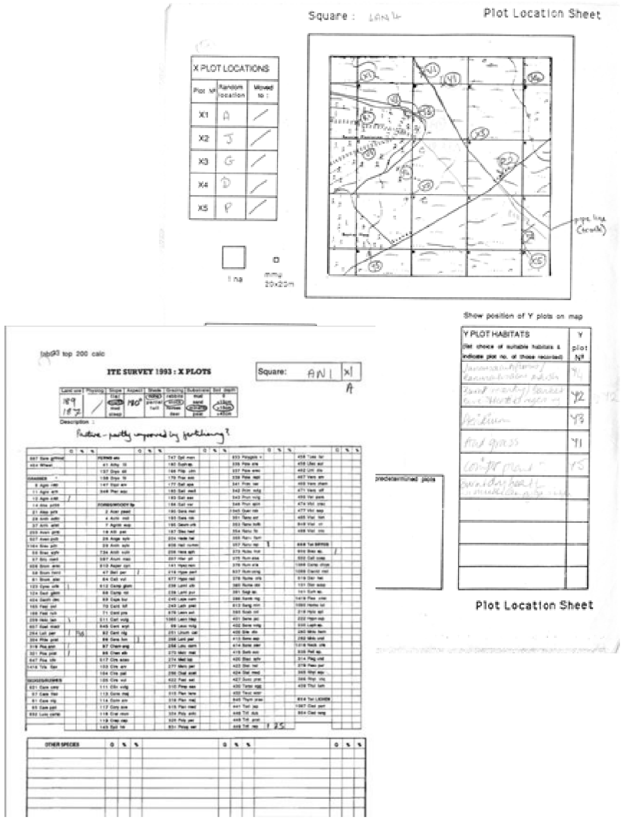
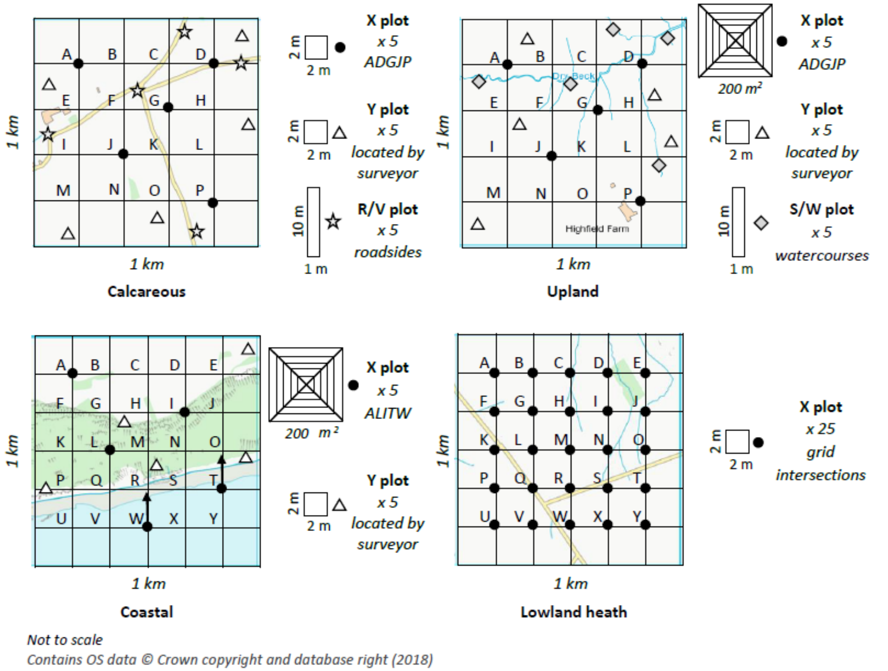
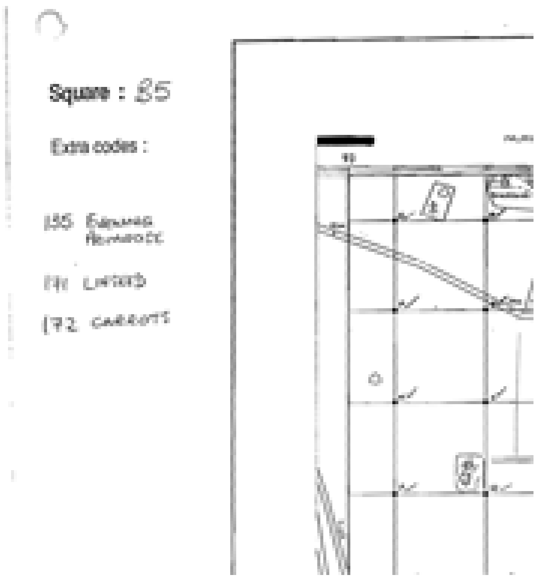


Figure ii.3 Paper 6 resources

Table ii.3. Data tasks (specific to the Paper 6 data):

Quality/data issue	Example	Correction/Solution	Script Ref.
Strata files			
Identify correct data from existing flat files (data sets of the same type were stored in separate Oracle tables in different schemas).	4 flat database files storing information on each of the 'Key Habitat' strata types: Calcareous, Upland Heath/Lowland heath/Coastal in different schema. • *LCOVER_COAST.csv* • *LCOVER_CHALKLIME_ENGL2.csv* • *LCOVER_LCLASS_HEATH_ENGL.csv* • *LCOVER_LCLASS_UPL_ENGL.csv* • 	Integrate data sets into single data set for efficiency and create data sets in spatial format.  Identify filter criteria from reports. Filter files to create spatial data then overlay to result in a final strata file. For example, in the coastal file, the data had to be filtered to result in the correct number of rows using Filter on Urban = o and Rejected = 0	KHS1

Quality/data issue	Example	Correction/Solution	Script Ref.
Plots			
No locations recorded with plot data	Data in the 'Key Habitat' survey were recorded at specified grid positions which had never been digitised. 	Identify the protocol for recording plots in each habitat type. Generate grids across sites, of the appropriate size for that type. Match intersections with plot types. Digitise and append additional plots that were off grid. Check field sheets for plots still missing location information. 	KHS2
Surveyor created codes not consistent	Surveyors often inserted square specific codes onto field sheets, inventing their own code list. This meant codes would overlap,	Create a lookup table of all surveyor specific codes. Update the numeric codes in the data to reflect this lookup table.	KHS3

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr>
<td></td>
<td>with a specific numeric number meaning different things at different sites.


</td>
<td>
<table>
<tr><th>SERIES_NU...</th><th>GRIDCODE</th><th>HABT</th><th>HABT_DESC</th><th>LUSE</th></tr>
<tr><td>B5</td><td>A</td><td>123</td><td>Potatoes</td><td>AC</td></tr>
<tr><td>B5</td><td>B</td><td>129</td><td>Other crop</td><td>AC</td></tr>
<tr><td>B5</td><td>C</td><td>129</td><td>Evening primro...</td><td>AC</td></tr>
<tr><td>B5</td><td>D</td><td>129</td><td>Evening primro...</td><td>AC</td></tr>
<tr><td>B5</td><td>E</td><td>120</td><td>Sugar beat</td><td>AC</td></tr>
<tr><td>B5</td><td>F</td><td>123</td><td>Potatoes</td><td>AC</td></tr>
<tr><td>B5</td><td>G</td><td>118</td><td>Barley</td><td>AC</td></tr>
<tr><td>B5</td><td>H</td><td>129</td><td>Evening primro...</td><td>AC</td></tr>
<tr><td>B5</td><td>I</td><td>120</td><td>Sugar beat</td><td>AC</td></tr>
<tr><td>B5</td><td>J</td><td>120</td><td>Sugar beat</td><td>AC</td></tr>
<tr><td>B5</td><td>K</td><td>123</td><td>Potatoes</td><td>AC</td></tr>
<tr><td>B5</td><td>L</td><td>118</td><td>Barley</td><td>AC</td></tr>
<tr><td>B5</td><td>M</td><td>129</td><td>Carrots</td><td>AC</td></tr>
<tr><td>B5</td><td>N</td><td>129</td><td>Linseed</td><td>AC</td></tr>
<tr><td>B5</td><td>O</td><td>129</td><td>Other crop</td><td>AC</td></tr>
<tr><td>B5</td><td>P</td><td>118</td><td>Barley</td><td>AC</td></tr>
<tr><td>B5</td><td>Q</td><td>118</td><td>Barley</td><td>AC</td></tr>
<tr><td>B5</td><td>R</td><td>118</td><td>Barley</td><td>AC</td></tr>
<tr><td>B5</td><td>S</td><td>129</td><td>Linseed</td><td>AC</td></tr>
<tr><td>B5</td><td>T</td><td>120</td><td>Sugar beat</td><td>AC</td></tr>
<tr><td>B5</td><td>U</td><td>118</td><td>Barley</td><td>AC</td></tr>
<tr><td>B5</td><td>V</td><td>118</td><td>Barley</td><td>AC</td></tr>
</table>
</td>
<td></td>
</tr>
</table>

Quality/data issue	Example	Correction/Solution	Script Ref.

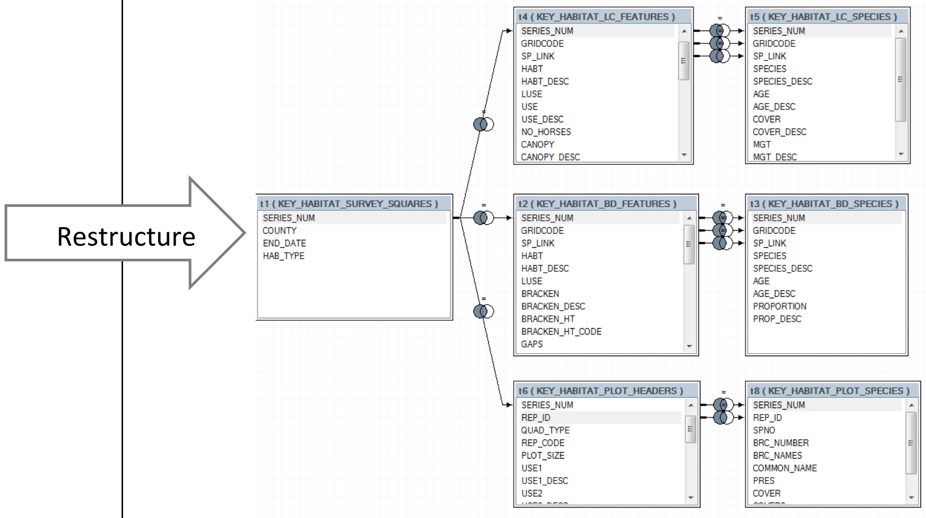
SQR	MAP	CODE	DESCRIPTION
E11	BD	249	ALNUS CORDATA
E12	LC	198	RECENTLY CATTLE GRAZED
E12	LC	135	MIX- HOL.LAN.,DES.CESP.,FEST.ARUN.,FEST.RUB.,AGRO.STOL.
B5	LC	135	EVENING PRIMROSE
B5	LC	171	LINSEED
B5	LC	172	CARROTS
B6	BD	324	PRUNUS SPINOSA
B6	BD	314	WIRE ON TRUNKS
B6	LC	135	LINUM FLAX LINSEED
B7	LC	172	CHAMERION ANGUSTIFOLIUM
B7	LC	171	GALIUM SAXATILE
B7	LC	135	LINSEED
B7	LC	173	CARROTS
B8	LC	171	FRENCH BEANS
B8	LC	172	RUMEX ACETOSELLA
B8	LC	198	RABBITS
B8	LC	173	URTICA DIOICA
B8	LC	135	TURF
E1	LC	171	RUMEX ACCLOSELLA
E1	LC	249	RHODODENDRON
E1	BD	249	RHODODENDRON
E1	BD	171	RUMEX ACCLOSELLA
E2	LC	171	LOTUS CORNICULATIS
E2	LC	434	W.D. TRAINING

Quality/data issue	Example	Correction/Solution	Script Ref.
Some plot data have negative cover values	Tree canopy covers have been recorded as negative values	Update negative values to positive, as it is standard Countryside Survey protocol to count overhanging trees species in the plot cover estimates.	N/A

SQUARE	REP	SPNO	PRES	COVER	COVER2	NAME
BD19	Y3	588	1	-100	0	Aesculus hippocastanum
BD15	Y3	162	1	-100	0	Fagus sylvatica
BN10	Y3	162	0	-100	0	Fagus sylvatica
BD16	Y2	170	1	-100	0	Fraxinus excelsior
BD14	Y1	117	0	-95	0	Corylus avellana
BD11	Y4	2	0	-95	0	Acer pseudoplatanus
BN10	V2	2	1	-90	0	Acer pseudoplatanus
BD14	Y2	170	0	-90	0	Fraxinus excelsior
BD13	Y2	19	1	-90	0	Alnus glutinosa
BD15	X1	162	0	-85	0	Fagus sylvatica
BD18	X3	384	0	-80	0	Salix cinerea
BD13	Y3	19	1	-80	0	Alnus glutinosa
BD1	Y4	596	0	-80	0	Populus nigra
BN5	Y2	678	1	-80	0	Tilia hybrids
BD17	Y5	1	0	-75	0	Acer campestre
BN10	X5	2	0	-75	0	Acer pseudoplatanus
BD17	Y5	170	1	-70	0	Fraxinus excelsior
AD5	R1	170	2	-70	0	Fraxinus excelsior
BD11	Y5	162	0	-70	0	Fagus sylvatica
BD22	X2	170	0	-65	0	Fraxinus excelsior
BN10	R2	685	2	-65	0	Tilia cordata
BD11	V1	344	0	-65	0	Prunus avium

SERIES_NU...	REP_ID	SPNO	BRC_NAMES	PRES	COVER	COVER2
BD19	Y3	204	Hedera helix	1	0	0
BD19	Y3	373	Rubus fruticosus agg.	1	0	0
BD19	Y3	462	Urtica dioica	1	0	0
BD19	Y3	588	Aesculus hippocastanum	1	100	0
BD19	Y3	887	Bare ground/rock	1	90	0
BD19	Y4	8	Elytrigia repens	1	5	0
BD19	Y4	11	Agrostis stolonifera	1	5	0
BD19	Y4	29	Anthriscus sylvestris	1	0	0
BD19	Y4	37	Arrhenatherum elatius	1	15	0
BD19	Y4	54	Brachythecium rutabulum	1	0	0
BD19	Y4	66	Calystegia sepium	1	0	0
BD19	Y4	92	Centaurea nigra	1	15	0
BD19	Y4	105	Cirsium vulgare	1	0	0
BD19	Y4	114	Convolvulus arvensis	1	0	0
BD19	Y4	118	Crataegus monogyna	1	5	0
BD19	Y4	124	Dactylis glomerata	1	5	0
BD19	Y4	161	Eurhynchium sp.	1	0	0
BD19	Y4	166	Festuca rubra agg.	1	15	0
BD19	Y4	177	Galium aparine	1	0	0
BD19	Y4	195	Geum urbanum	1	0	0
BD19	Y4	197	Glechoma hederacea	1	0	0
BD19	Y4	249	Ligustrum vulgare	1	10	0
BD19	Y4	304	Phleum pratense sens.lat.	1	0	0
BD19	Y4	315	Plantago lanceolata	1	10	0
BD19	Y4	346	Prunus spinosa	1	5	0

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr><td>Features:</td><td></td><td></td><td></td></tr>
<tr><td>Find any available extant data in old Oracle accounts (LUST). Each table named and structured differently according to landscape type.</td><td>COAST_FABLC.csv
FABLC_CALC_PRIM.csv
FABLC_CALC93_PRIM.csv
FABLC_UPL.csv
FABLC_UPL93_PRIM.csv
HEATH_PRIM_LC.csv
HEATH_TEXT.csv

FA_COVER_1992.csv
FA_COVER_1993.csv
KEY_FABS2.csv
KEY_FABS3.csv
COAST_FABBD93.csv
FABBD_CALC_PRIM.csv
FABBD_CALC93_PRIM.csv
HEATH_BOUNDARIES.csv</td><td>
</td><td>N/A</td></tr>
<tr><td>Land Cover: data located in old Oracle databases were stored with attribute codes in separate rows, thus unrelated to the primary feature.</td><td><table>
<tr><th>GRIDCO</th><th>FAB_PO</th><th>HAB_TYPE</th><th>CODE</th><th>DESCRIPTION</th><th>ADD_CC</th><th>ADD_DESCRIPTION</th></tr>
<tr><td>A</td><td>1</td><td>Lowland Heath</td><td>107</td><td>Lowland heath</td><td></td><td></td></tr>
<tr><td>A</td><td>2</td><td>Lowland Heath</td><td>171</td><td>AGRO CURTISII</td><td>171</td><td>AGRO CURTISII</td></tr>
<tr><td>A</td><td>3</td><td>Lowland Heath</td><td>176</td><td>50-75%</td><td></td><td></td></tr>
<tr><td>A</td><td>4</td><td>Lowland Heath</td><td>157</td><td>Pteridium aquilinum - scattere</td><td></td><td></td></tr>
<tr><td>A</td><td>5</td><td>Lowland Heath</td><td>175</td><td>25-50%</td><td></td><td></td></tr>
<tr><td>A</td><td>6</td><td>Lowland Heath</td><td>183</td><td>1-1.5m</td><td></td><td></td></tr>
<tr><td>A</td><td>7</td><td>Lowland Heath</td><td>239</td><td>Gorse</td><td></td><td></td></tr>
<tr><td>A</td><td>8</td><td>Lowland Heath</td><td>175</td><td>25-50%</td><td></td><td></td></tr>
<tr><td>A</td><td>9</td><td>Lowland Heath</td><td>191</td><td>Horses</td><td></td><td></td></tr>
<tr><td>A</td><td>10</td><td>Lowland Heath</td><td>0</td><td></td><td></td><td></td></tr>
<tr><td>A</td><td>11</td><td>Lowland Heath</td><td>195</td><td>Deer</td><td></td><td></td></tr>
<tr><td>B</td><td>1</td><td>Lowland Heath</td><td>107</td><td>Lowland heath</td><td></td><td></td></tr>
<tr><td>B</td><td>2</td><td>Lowland Heath</td><td>161</td><td>Calluna vulgaris</td><td></td><td></td></tr>
</table>
In this example, in the first row, the primary feature is 107, Lowland Heath. Then 171 (Agrostis curtisii) is an attribute of the Lowland heath, and 176 (50-75%) refers to the coverage of the Agrostis curtisii. However, once the table is reordered, this relationship is immediately lost.</td><td>Data had to be carefully transposed to associate the correct codes with the correct features. A relationship had to be created between items recorded against the same attribute, based on their position in the table.</td><td>KHS4, KHS8</td></tr>
</table>

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr><td></td><td></td><td>
<table>
<tr><th>SERIES_NU...</th><th>GRIDCODE</th><th>SPECIES</th><th>SPECIES_DESC</th><th>AGE</th><th>AGE_DE...</th><th>COVER</th><th>COVER_DE...</th></tr>
<tr><td>A1</td><td>A</td><td>157</td><td>Pteridium aquilinum - scattered</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>A</td><td>239</td><td>Gorse</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>B</td><td>161</td><td>Calluna vulgaris</td><td>.</td><td></td><td>178</td><td>95-100%</td></tr>
<tr><td>A1</td><td>C</td><td>156</td><td>Pteridium aquilinum - dense</td><td>.</td><td></td><td>177</td><td>75-95%</td></tr>
<tr><td>A1</td><td>C</td><td>163</td><td>Molinia caerulea</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>C</td><td>171</td><td>Agrostis curtisii</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>C</td><td>174</td><td>Agrostis canina</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>D</td><td>161</td><td>Calluna vulgaris</td><td>.</td><td></td><td>178</td><td>95-100%</td></tr>
<tr><td>A1</td><td>E</td><td>167</td><td>Sphagnum spp.</td><td>.</td><td></td><td>176</td><td>50-75%</td></tr>
<tr><td>A1</td><td>E</td><td>163</td><td>Molinia caerulea</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>E</td><td>173</td><td>Narthecium ossifragum</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>F</td><td>163</td><td>Molinia caerulea</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>F</td><td>169</td><td>Juncus squarrosus</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>F</td><td>161</td><td>Calluna vulgaris</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>A1</td><td>G</td><td>161</td><td>Calluna vulgaris</td><td>.</td><td></td><td>177</td><td>75-95%</td></tr>
<tr><td>A1</td><td>G</td><td>163</td><td>Molinia caerulea</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>H</td><td>163</td><td>Molinia caerulea</td><td>.</td><td></td><td>176</td><td>50-75%</td></tr>
<tr><td>A1</td><td>H</td><td>169</td><td>Juncus squarrosus</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>H</td><td>173</td><td>Narthecium ossifragum</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>H</td><td>161</td><td>Calluna vulgaris</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>A1</td><td>I</td><td>161</td><td>Calluna vulgaris</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>I</td><td>169</td><td>Juncus squarrosus</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>I</td><td>163</td><td>Molinia caerulea</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>J</td><td>161</td><td>Calluna vulgaris</td><td>.</td><td></td><td>178</td><td>95-100%</td></tr>
<tr><td>A1</td><td>J</td><td>157</td><td>Pteridium aquilinum - scattered</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>A1</td><td>K</td><td>163</td><td>Molinia caerulea</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>K</td><td>169</td><td>Juncus squarrosus</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>K</td><td>161</td><td>Calluna vulgaris</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>A1</td><td>L</td><td>169</td><td>Juncus squarrosus</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
<tr><td>A1</td><td>L</td><td>163</td><td>Molinia caerulea</td><td>.</td><td></td><td>175</td><td>25-50%</td></tr>
</table>
<table>
<tr><th>HABT</th><th>HABT_DESC</th><th>LUSE</th><th>USE</th><th>USE_DESC</th><th>NO_HORSES</th><th>CAN</th></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>195</td><td>Deer</td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>4</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>4</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>2</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>191</td><td>Horses</td><td>0</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
<tr><td>107</td><td>Lowland heath</td><td>AN</td><td>.</td><td></td><td>.</td><td></td></tr>
</table>
</td><td></td></tr>
<tr><td>Land cover species data</td><td>Species have a many to one relationship with the recording location, hence need to be stored in a separate table to the primary codes, along with the % cover, age, management and</td><td>Select species from data and create new species attribute table.</td><td>KHS5</td></tr>
</table>

Quality/data issue	Example	Correction/Solution	Script Ref.
	height information.		

GRIDCODE	FAB_POSITION	HAB_TYPE	CODE	DESCRIPTION	ADD_CODE	ADD_DESCRIPTION
A	1	Lowland Heath	107	Lowland heath		
A	2	Lowland Heath	171	AGRO CURTISII	171	AGRO CURTISII
A	3	Lowland Heath	176	50-75%		
A	4	Lowland Heath	157	Pteridium aquilinum - scattere		
A	5	Lowland Heath	175	25-50%		
A	6	Lowland Heath	183	1-1.5m		
A	7	Lowland Heath	239	Gorse		
A	8	Lowland Heath	175	25-50%		
A	9	Lowland Heath	191	Horses		
A	10	Lowland Heath	0			
A	11	Lowland Heath	195	Deer		
B	1	Lowland Heath	107	Lowland heath		
B	2	Lowland Heath	161	Calluna vulgaris		

SERIES_NUM	GRIDCODE	SPECIES	SPECIES_DESC	AGE	AGE_DESC	COVER	COVER_DESC	MGT	MGT_DESC	HEIGHT	HEIGHT_
A1	R	172	Cladonia portentosa			175	25-50%				
A1	S	161	Calluna vulgaris			176	50-75%			179	< 10cm
A1	T	223	Pine - corsican			256	25-50%				
A1	T	234	Birch	262	5-20 years	256	25-50%	276	Unmanaged - thriving		
A1	T	295	Bracken - scattered			258	75-95%	282	Natural regeneration		
A1	U	161	Calluna vulgaris			177	75-95%			181	30-50cm
A1	V	163	Molinia caerulea			176	50-75%				
A1	V	169	Juncus squarrosus			176	50-75%				
A1	V	161	Calluna vulgaris							181	30-50cm
A1	W	156	Pteridium aquilinum - dense			177	75-95%			183	1-1.5m
A1	W	163	Molinia caerulea			256	25-50%				
A1	W	161	Calluna vulgaris							181	30-50cm
A1	X	225	Pine - scots	263	20-100 years	258	75-95%	275	Well managed		
A1	X	294	Bracken - dense			258	75-95%				
A1	Y	233	Beech			257	50-75%				
A1	Y	225	Pine - scots	263	20-100 years	256	25-50%	277	Unmanaged - improvable		
A1	Y	295	Bracken - scattered								

Quality/data issue	Example	Correction/Solution	Script Ref.
Boundaries: data located in old Oracle databases were stored with attribute codes in separate rows, thus unrelated to the primary feature.		Data had to be carefully related to associate the correct codes with the correct features. A relationship had to be created between items recorded against the same attribute, based on their position in the table.	KHS6, KHS8

	A	B	C	D	E	F	G	H
1	Idno	SERIES_NUM	CODE_TYPE	GRIDCODE	FAB_POSITION	HAB_TYPE	CODE	DESCRIPTION
2	26	A1	BD	J	3	Lowland Heath	351	Stockproof
3	27	A1	BD	O	3	Lowland Heath	351	Stockproof
4	3	A1	BD	O	6	Lowland Heath	215	Closed canopy
5	28	A1	BD	S	3	Lowland Heath	351	Stockproof
6	29	A1	BD	T	3	Lowland Heath	351	Stockproof
7	30	A1	BD	W	3	Lowland Heath	351	Stockproof
8	31	A1	BD	X	3	Lowland Heath	351	Stockproof
9	265	A10	BD	A	3	Lowland Heath	352	Not stockproof
10	266	A10	BD	D	3	Lowland Heath	352	Not stockproof
11	267	A10	BD	G	3	Lowland Heath	352	Not stockproof
12	268	A10	BD	L	3	Lowland Heath	352	Not stockproof
13	269	A10	BD	Q	3	Lowland Heath	352	Not stockproof
14	270	A10	BD	U	3	Lowland Heath	352	Not stockproof
15	271	A10	BD	W	3	Lowland Heath	352	Not stockproof
16	272	A10	BD	W	4	Lowland Heath	359	Derelict
17	273	A10	BD	W	5	Lowland Heath	363	Regrowth from stumps
18	274	A10	BD	W	6	Lowland Heath	364	Bracken present
19	598	A11	BD	C	3	Lowland Heath	352	Not stockproof
20	606	A11	BD	C	4	Lowland Heath	359	Derelict
21	627	A11	BD	C	7	Lowland Heath	364	Bracken present
22	599	A11	BD	E	3	Lowland Heath	352	Not stockproof
23	607	A11	BD	E	4	Lowland Heath	359	Derelict
24	620	A11	BD	E	5	Lowland Heath	364	Bracken present

Sq	Grid	HABT	HABT_DESC	LUSE	BRACKEN	BRACKEN_DESC	BRACKEN_HT_CODE	BRACKEN_HT	GAPS	GAPS_DESC	MGT	MGT_DESC	STOCK	STOCK_DESC
A1	J	313	Fence - wire on posts	F									351	Stockproof
A1	O	57	Roadside ditch	IW	295	Bracken - scattered							351	Stockproof
A1	O	203	Line of trees	FO	295	Bracken - scattered							351	Stockproof
A1	O	313	Fence - wire on posts	F	295	Bracken - scattered							351	Stockproof
A1	S	313	Fence - wire on posts	F									351	Stockproof
A1	T	313	Fence - wire on posts	F									351	Stockproof
A1	W	313	Fence - wire on posts	F									351	Stockproof
A1	X	313	Fence - wire on posts	F									351	Stockproof
A10	A	313	Fence - wire on posts	F									352	Not stockproo
A10	D	311	Fence - wood only	F									352	Not stockproo
A10	G	313	Fence - wire on posts	F									352	Not stockproo
A10	L	313	Fence - wire on posts	F									352	Not stockproo
A10	Q	313	Fence - wire on posts	F									352	Not stockproo
A10	U	313	Fence - wire on posts	F									352	Not stockproo
A10	W	323	Mixed hedge	WLF	364	Bracken present					359	Derelict	352	Not stockproo
A11	C	313	Fence - wire on posts	F	364	Bracken present	183	1-1.5m			359	Derelict	352	Not stockproo
A11	C	332	Earth bank	B	364	Bracken present	183	1-1.5m			359	Derelict	352	Not stockproo
A11	E	313	Fence - wire on posts	F	364	Bracken present	183	1-1.5m			359	Derelict	352	Not stockproo
A11	G	313	Fence - wire on posts	F	364	Bracken present	183	1-1.5m			359	Derelict	352	Not stockproo
A11	G	332	Earth bank	B	364	Bracken present	183	1-1.5m			359	Derelict	352	Not stockproo
A11	I	313	Fence - wire on posts	F	364	Bracken present	183	1-1.5m			359	Derelict	352	Not stockproo
A11	J	313	Fence - wire on posts	F	364	Bracken present	183	1-1.5m			359	Derelict	352	Not stockproo

Quality/data issue	Example	Correction/Solution	Script Ref.
Boundary related species data	As with the land cover data, species have a many to one relationship with the recording location, hence need to be stored in a separate table to the primary codes, along with the % cover and age.	Data was restructured as below:	KHS7

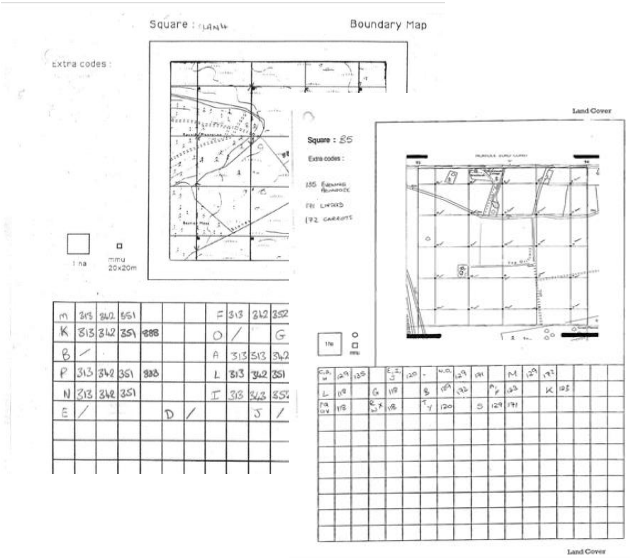
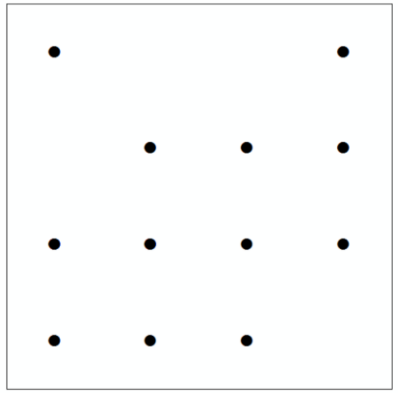
Quality/data issue	Example	Correction/Solution	Script Ref.
	(see spreadsheet extract below)	(see spreadsheet extract below)	

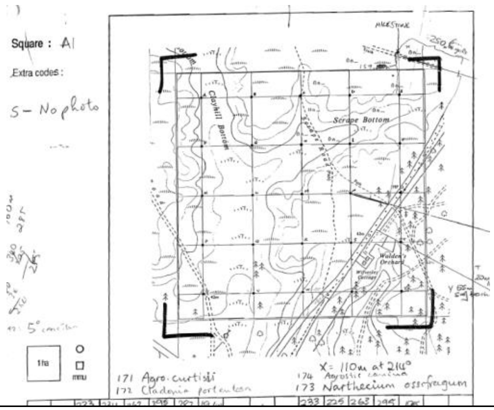
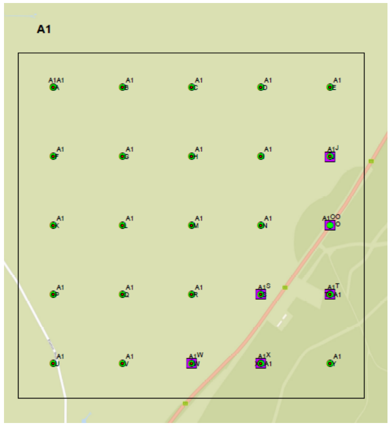
Example

	A	B	C	D	E	F	G	
1	Idno	SERIES_NUM	CODE_TYPE	GRIDCODE	FAB_POSITION	HAB_TYPE	CODE	DESCRIPTION
2	26	A1	BD	J	3	Lowland Heath	351	Stockproof
3	27	A1	BD	O	3	Lowland Heath	351	Stockproof
4	3	A1	BD	O	6	Lowland Heath	215	Closed canopy
5	28	A1	BD	S	3	Lowland Heath	351	Stockproof
6	29	A1	BD	T	3	Lowland Heath	351	Stockproof
7	30	A1	BD	W	3	Lowland Heath	351	Stockproof
8	31	A1	BD	X	3	Lowland Heath	351	Stockproof
9	265	A10	BD	A	3	Lowland Heath	352	Not stockproof
10	266	A10	BD	D	3	Lowland Heath	352	Not stockproof
11	267	A10	BD	G	3	Lowland Heath	352	Not stockproof
12	268	A10	BD	L	3	Lowland Heath	352	Not stockproof
13	269	A10	BD	Q	3	Lowland Heath	352	Not stockproof
14	270	A10	BD	U	3	Lowland Heath	352	Not stockproof
15	271	A10	BD	W	3	Lowland Heath	352	Not stockproof
16	272	A10	BD	W	4	Lowland Heath	359	Derelict
17	273	A10	BD	W	5	Lowland Heath	363	Regrowth from stumps
18	274	A10	BD	W	6	Lowland Heath	364	Bracken present
19	598	A11	BD	C	3	Lowland Heath	352	Not stockproof
20	606	A11	BD	C	4	Lowland Heath	359	Derelict
21	627	A11	BD	C	7	Lowland Heath	364	Bracken present
22	599	A11	BD	E	3	Lowland Heath	352	Not stockproof
23	607	A11	BD	E	4	Lowland Heath	359	Derelict
24	620	A11	BD	E	5	Lowland Heath	364	Bracken present

Correction/Solution

	A	B	C	D	E	F	G	H	I
1	Sq	Grid	Group3	SPECIES	SPECIES_DESC	AGE	AGE_DESC	PROPORTION	PROP_DESC
2	A1	O	1	243	Oak	263	20-100 years	256	25-50%
3	A1	O	2	233	Beech	263	20-100 years	256	25-50%
4	A1	O	3	246	Sweet chestnut	263	20-100 years	256	25-50%
5	A11	V	4	249	Ilex aquifolium			259	95-100%
6	A4	F	6	243	Oak	264	> 100 years		
7	A4	F	7	324	Hazel				
8	A4	G	8	248	Willow				
9	A4	G	9	243	Oak	264	> 100 years		
10	A4	U	10	243	Oak				
11	A4	U	11	232	Ash	264	> 100 years		
12	A4	V	12	324	Hazel				
13	A4	X	13	237	Elm	262	5-20 years		
14	A4	X	14	235	Bramble				
15	A5	G	15	242	Lime	263	20-100 years		
16	A5	H	16	243	Oak				
17	A5	L	17	243	Oak	263	20-100 years		
18	A5	L	18	232	Ash	263	20-100 years		
19	A5	X	19	235	Bramble				
20	A5	Y	20	235	Bramble				
21	A6	E	21	225	Pine - scots	264	> 100 years	259	95-100%

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr><td>Recording locations missing from data</td><td>Locate paper maps
</td><td>Identify the protocol for recording plots in each habitat type. Generate grids across sites, of the appropriate size for that type. Match intersections with recording locations, as for the plot example, above.

Example of recording locations in a survey square</td><td>N/A</td></tr>
<tr><td>Attribute codes with a prefix of “– “ indicate counts of items (e.g. number of horses).</td><td><table>
<tr><td>LC</td><td>F</td><td>5</td><td>Lowland Heath</td><td>175</td><td>25-50%</td></tr>
<tr><td>LC</td><td>F</td><td>6</td><td>Lowland Heath</td><td>161</td><td>Calluna vulgaris</td></tr>
<tr><td>LC</td><td>F</td><td>7</td><td>Lowland Heath</td><td>180</td><td>10-30cm</td></tr>
<tr><td>LC</td><td>F</td><td>8</td><td>Lowland Heath</td><td>91</td><td>Regenerating heather</td></tr>
<tr><td>LC</td><td>F</td><td>9</td><td>Lowland Heath</td><td>191</td><td>Horses</td></tr>
<tr><td>LC</td><td>F</td><td>10</td><td>Lowland Heath</td><td>-4</td><td></td></tr>
<tr><td>LC</td><td>G</td><td>1</td><td>Lowland Heath</td><td>107</td><td>Lowland heath</td></tr>
<tr><td>LC</td><td>G</td><td>2</td><td>Lowland Heath</td><td>161</td><td>Calluna vulgaris</td></tr>
<tr><td>LC</td><td>G</td><td>3</td><td>Lowland Heath</td><td>177</td><td>75-95%</td></tr>
<tr><td>LC</td><td>G</td><td>4</td><td>Lowland Heath</td><td>182</td><td>0.5-1m</td></tr>
</table></td><td>Create an additional field named ‘NO_HORSES’ and populate with values, transforming them into positive values, rather than negative.</td><td>KHS4</td></tr>
</table>

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr><td></td><td></td><td>
<table>
<tr><th>SERIES_NU...</th><th>GRIDCODE</th><th>HABT</th><th>HABT_DESC</th><th>NO_HORSES</th><th>LUSE</th></tr>
<tr><td>G14</td><td>M</td><td>206</td><td>Woodland/forest</td><td>.</td><td>FO</td></tr>
<tr><td>G14</td><td>M</td><td>404</td><td>Public open sp...</td><td>.</td><td>RE</td></tr>
<tr><td>G14</td><td>N</td><td>402</td><td>Garden/ground...</td><td>.</td><td>ST</td></tr>
<tr><td>G14</td><td>O</td><td>206</td><td>Woodland/forest</td><td>.</td><td>FO</td></tr>
<tr><td>G14</td><td>P</td><td>101</td><td>Lowland agricul...</td><td>7</td><td>AN</td></tr>
<tr><td>G14</td><td>Q</td><td>403</td><td>Garden/ground...</td><td>.</td><td>ST</td></tr>
<tr><td>G14</td><td>R</td><td>206</td><td>Woodland/forest</td><td>.</td><td>FO</td></tr>
<tr><td>G14</td><td>R</td><td>404</td><td>Public open sp...</td><td>.</td><td>RE</td></tr>
<tr><td>G14</td><td>S</td><td>457</td><td>Unconstructed...</td><td>.</td><td>TR</td></tr>
<tr><td>EN4</td><td>D</td><td>101</td><td>Lowland agricul...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>E</td><td>114</td><td>Marsh</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>F</td><td>106</td><td>Maritime veget...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>F</td><td>608</td><td>Maritime grassl...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>G</td><td>106</td><td>Maritime veget...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>G</td><td>608</td><td>Maritime grassl...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>H</td><td>106</td><td>Maritime veget...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>H</td><td>608</td><td>Maritime grassl...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>I</td><td>106</td><td>Maritime veget...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>I</td><td>608</td><td>Maritime grassl...</td><td>.</td><td>AN</td></tr>
<tr><td>EN4</td><td>J</td><td>53</td><td>River</td><td>.</td><td>IW</td></tr>
</table>
</td><td></td></tr>
<tr><td>Checks against recording sheets</td><td>In order to ensure data quality and identify and consistent sources of error (particularly in the case of recoding locations), undertake random spot-checks comparing final data sets with field sheets to identify any systematic errors. Reiterate processes as necessary.

</td><td>
</td><td>N/A</td></tr>
</table>

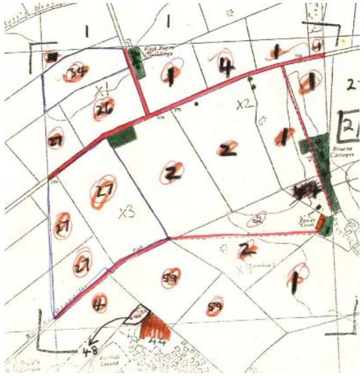
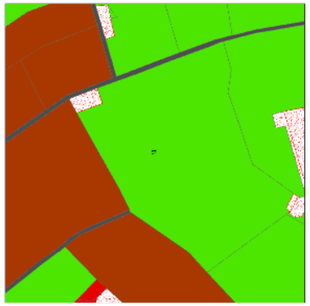
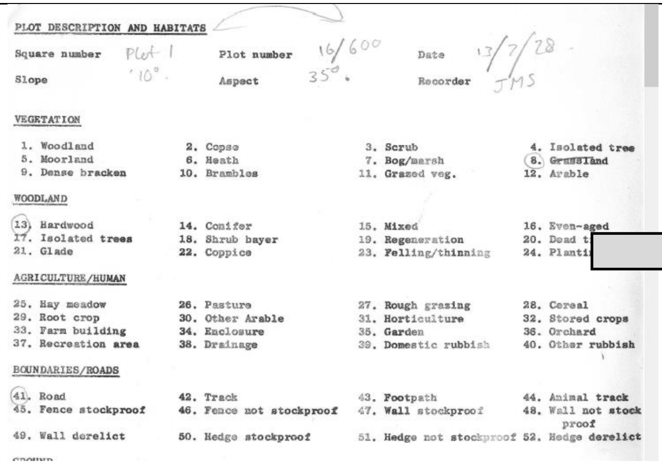
Countryside Survey (papers 4 and 5)

Resources available (specific to the Paper 4 and 5 data):

- *Field handbooks*
- *Majority of data securely stored in Oracle databases with the exception of those described below*
- *Reports*
- *Website*
- *Published papers*

Table ii.4. Data tasks (specific to the Paper 4 and 5 data):

Quality/data issue	Example	Correction/Solution	Script Ref.
Landscape mapping data missing from 1978	The 256 annotated land-cover maps for the 1978 Countryside Survey squares had never been digitised into a GIS format before 2009 (although the areas and lengths had been recorded using rudimentary digitising tools in the 1980s). The field handbook from 1978 has a paragraph explaining the methodology for mapping habitats across survey squares. The first few habitat maps were hand drawn sketches on blank paper, then later transferred onto Ordnance Survey 1:10 000 base maps. The sketches and field maps were transferred onto the base maps using a set of 80 codes which were mainly species descriptions but were based on traditionally taught divisions which, in most cases, have helped them translate easily to Broad Habitats (Jackson, 2000) and in some cases, Priority Habitats (Maddock, 2008).	A UKCEH Data Rescue project undertaken in 2009 allowed digitisation of the maps (Wood, 2012). • Firstly, the maps were scanned on an auto-feed scanner at CEH Lancaster and are available in .jpg format and .pdf format. • Plot data sheets were also scanned and placed with other CS data sheets. • The 256 maps of habitat areas were then digitized by ADAS in summer 2009, according to a defined protocol set out in Appendix ii of CS1978 – Data Rescue Scoping study for digitizing Countryside Survey primary field data documents, Final Report (Wood, 2008). • Survey used unique set of codes translated into Broad Habitat codes	*See Wood et al.,2012 for further details.* Digitising protocol (CS9).

<table>
<tr><th>Quality/data issue</th><th>Example</th><th>Correction/Solution</th><th>Script Ref.</th></tr>
<tr><td></td><td></td><td></td><td></td></tr>
<tr><td>Plot attribute data missing from 1978</td><td>
PLOT DESCRIPTION AND HABITATS
Square number Plot 1 Plot number 16/600 Date 13/7/28
Slope 10° Aspect 35° Recorder JMS

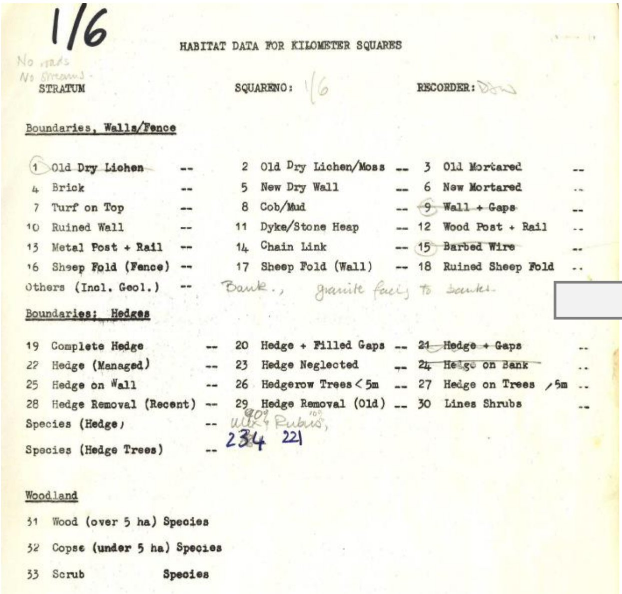
VEGETATION
1. Woodland 2. Copse 3. Scrub 4. Isolated tree
5. Moorland 6. Heath 7. Bog/marsh 8. Grassland
9. Dense bracken 10. Brambles 11. Grazed veg. 12. Arable

WOODLAND
13. Hardwood 14. Conifer 15. Mixed 16. Even-aged
17. Isolated trees 18. Shrub bayer 19. Regeneration 20. Dead t
21. Glade 22. Coppice 23. Felling/thinning 24. Planti

AGRICULTURE/HUMAN
25. Hay meadow 26. Pasture 27. Rough grazing 28. Cereal
29. Root crop 30. Other Arable 31. Horticulture 32. Stored crops
33. Farm building 34. Enclosure 35. Garden 36. Orchard
37. Recreation area 38. Drainage 39. Domestic rubbish 40. Other rubbish

BOUNDARIES/ROADS
41. Road 42. Track 43. Footpath 44. Animal track
45. Fence stockproof 46. Fence not stockproof 47. Wall stockproof 48. Wall not stock proof
49. Wall derelict 50. Hedge stockproof 51. Hedge not stockproof 52. Hedge derelict</td><td>Data digitised by student and uploaded to Oracle database

<table>
<tr><th>SERIES_NU...</th><th>PLOT_TYPE</th><th>PLOT_NO</th><th>REP_ID</th><th>CODE</th><th>DESCRIPTION</th><th>CODE_GROUP</th></tr>
<tr><td>18</td><td>X</td><td>5</td><td>18X5</td><td>8</td><td>Grassland</td><td>VEGETATION</td></tr>
<tr><td>18</td><td>X</td><td>5</td><td>18X5</td><td>11</td><td>Grazed veg.</td><td>VEGETATION</td></tr>
<tr><td>18</td><td>X</td><td>5</td><td>18X5</td><td>26</td><td>Pasture</td><td>AGRICULTURE/HUMAN</td></tr>
<tr><td>18</td><td>X</td><td>5</td><td>18X5</td><td>50</td><td>Hedge stockproof</td><td>BOUNDARIES/ROADS</td></tr>
<tr><td>18</td><td>X</td><td>5</td><td>18X5</td><td>61</td><td>Bank</td><td>GROUND</td></tr>
<tr><td>18</td><td>X</td><td>5</td><td>18X5</td><td>82</td><td>Cattle</td><td>ANIMALS</td></tr>
<tr><td>18</td><td>X</td><td>5</td><td>18X5</td><td>96</td><td>Simple</td><td>TOPOGRAPHY</td></tr>
<tr><td>22</td><td>X</td><td>1</td><td>22X1</td><td>5</td><td>Moorland</td><td>VEGETATION</td></tr>
<tr><td>22</td><td>X</td><td>1</td><td>22X1</td><td>27</td><td>Rough grazing</td><td>AGRICULTURE/HUMAN</td></tr>
<tr><td>22</td><td>X</td><td>1</td><td>22X1</td><td>45</td><td>Fence stockproof</td><td>BOUNDARIES/ROADS</td></tr>
<tr><td>22</td><td>X</td><td>1</td><td>22X1</td><td>56</td><td>Boulders</td><td>GROUND</td></tr>
<tr><td>22</td><td>X</td><td>1</td><td>22X1</td><td>61</td><td>Bank</td><td>GROUND</td></tr>
<tr><td>22</td><td>X</td><td>1</td><td>22X1</td><td>92</td><td>Omni. birds</td><td>ANIMALS</td></tr>
<tr><td>22</td><td>X</td><td>1</td><td>22X1</td><td>95</td><td>Slightly variable</td><td>TOPOGRAPHY</td></tr>
<tr><td>22</td><td>X</td><td>1</td><td>22X1</td><td>96</td><td>Simple</td><td>TOPOGRAPHY</td></tr>
<tr><td>22</td><td>X</td><td>2</td><td>22X2</td><td>12</td><td>Arable</td><td>VEGETATION</td></tr>
<tr><td>22</td><td>X</td><td>2</td><td>22X2</td><td>28</td><td>Cereal</td><td>AGRICULTURE/HUMAN</td></tr>
<tr><td>22</td><td>X</td><td>2</td><td>22X2</td><td>42</td><td>Track</td><td>BOUNDARIES/ROADS</td></tr>
<tr><td>22</td><td>X</td><td>2</td><td>22X2</td><td>45</td><td>Fence stockproof</td><td>BOUNDARIES/ROADS</td></tr>
<tr><td>22</td><td>X</td><td>2</td><td>22X2</td><td>61</td><td>Bank</td><td>GROUND</td></tr>
<tr><td>22</td><td>X</td><td>2</td><td>22X2</td><td>90</td><td>Game birds</td><td>ANIMALS</td></tr>
<tr><td>22</td><td>X</td><td>2</td><td>22X2</td><td>96</td><td>Simple</td><td>TOPOGRAPHY</td></tr>
<tr><td>22</td><td>X</td><td>3</td><td>22X3</td><td>8</td><td>Grassland</td><td>VEGETATION</td></tr>
<tr><td>22</td><td>X</td><td>3</td><td>22X3</td><td>11</td><td>Grazed veg.</td><td>VEGETATION</td></tr>
<tr><td>22</td><td>X</td><td>3</td><td>22X3</td><td>26</td><td>Pasture</td><td>AGRICULTURE/HUMAN</td></tr>
<tr><td>22</td><td>X</td><td>3</td><td>22X3</td><td>45</td><td>Fence stockproof</td><td>BOUNDARIES/ROADS</td></tr>
</table></td><td>N/A</td></tr>
</table>

Quality/data issue	Example	Correction/Solution	Script Ref.
Missing code descriptions from landscape mapping data		Lookup table missing – recreated lookup table by generating a list of all possible codes from entered data, then adding descriptions from data sheets. Data digitised and uploaded to Oracle database.	N/A

1/6

HABITAT DATA FOR KILOMETER SQUARES

No roads
No Streams -
STRATUM SQUARENO: 1/6 RECORDER: DJW

Boundaries, Walls/Fence

1 Old Dry Lichen -- 2 Old Dry Lichen/Moss -- 3 Old Mortared --
4 Brick -- 5 New Dry Wall -- 6 New Mortared --
7 Turf on Top -- 8 Cob/Mud -- 9 Wall + Gaps --
10 Ruined Wall -- 11 Dyke/Stone Heap -- 12 Wood Post + Rail --
13 Metal Post + Rail -- 14 Chain Link -- 15 Barbed Wire --
16 Sheep Fold (Fence) -- 17 Sheep Fold (Wall) -- 18 Ruined Sheep Fold --
Others (Incl. Geol.) -- Bank., granite facing to banks.

Boundaries: Hedges

19 Complete Hedge -- 20 Hedge + Filled Gaps -- 21 Hedge + Gaps --
22 Hedge (Managed) -- 23 Hedge Neglected -- 24 Hedge on Bank --
25 Hedge on Wall -- 26 Hedgerow Trees < 5m -- 27 Hedge on Trees > 5m --
28 Hedge Removal (Recent) -- 29 Hedge Removal (Old) -- 30 Lines Shrubs --
Species (Hedge) -- 90% Ulex, Rubus 10%
Species (Hedge Trees) -- 234 221

Woodland

31 Wood (over 5 ha) Species
32 Copse (under 5 ha) Species
33 Scrub Species

SERIES_NUM	THEME	PCODE	PDESC	SPCODE	SPECIES	BRC_SPECIES	NOTES
6	Boundaries: Hedges	21	Hedge and gaps				
6	Boundaries: Hedges	24	Hedge on bank				
6	Boundaries: Hedges	301	Species (hedge)	221	Bramble	Rubus sp.	
6	Boundaries: Hedges	301	Species (hedge)	234	Gorse	Ulex sp.	
6	Boundaries: Walls/Fences	15	Barbed wire				
6	Boundaries: Walls/Fences	1	Old dry lichen				
6	Boundaries: Walls/Fences		Other				Bank, granite facing to banks
6	Boundaries: Walls/Fences	9	Wall and gaps				
6	Buildings (amenity)	203	Footpath <1m				
6	Buildings (amenity)	204	Footpath >1m				
6	Buildings (amenity)	209	Track				
6	Buildings (domestic)	168	Detached house				
6	Buildings (domestic)	175	Pre 1800				
6	Buildings (domestic)	176	Slate roofs				
6	Buildings (industrial)		Other				old mine shafts, standing stones
6	Domestic animals	97	Cattle				
6	Domestic animals	97	Cattle	449	Hereford		
6	Domestic animals	97	Cattle	451	Jersey		
6	Domestic animals	97	Cattle	467	Freesian x Hereford		
6	Habitats (rock)	110	Boulders				
6	Habitats (vegetation)	50	Calluna				

Analyses (Paper 7)

Paper 7 demonstrates an example of the kind of simple, comprehensive analyses that can be undertaken with national ecological survey data. A comprehensive, definitive set of statistics for the British uplands is provided, using a formal structure, with landscapes at the highest level followed by habitats, then vegetation, and finally species, for Great Britain (GB) as a whole, and also England, Wales and Scotland. The analyses scripts are documented in scripts CS2-8.

- Broad habitats – extent in the uplands
- Grazing animals – extent in the uplands
- Boundaries – linear landscape features – extent in the uplands
- Top 10 Countryside Vegetation System classes in the uplands
- Top 10 species and % cover in the uplands

Appendix iii: Code and Scripts

This appendix contains examples of code and examples of methodology and scripts referred to in Chapter 3 and Appendix ii.
The majority of the code was implemented in SAS Enterprise Guide (www.sas.com) using SQL and SAS scripting. Code has been exported from SAS Enterprise Guide using the export function, with 'wrapper' text removed. Imported data refers to data as entered into MS Excel or comma separated values files in the data entry process.
Additional documents describing methods not documented elsewhere are also included.

Dataset/Paper	Script ID	Description
Data manipulation		
Paper 1 (Pinewoods)	***SPW1***	Code to transpose diameter at breast height (DBH) codes
	SPW2	Code to incorporate additional information into plot ground flora table
	SPW3	Code to transpose plot description codes (soil)
Paper 2 (Woodlands)	***GBW1***	Code to transform Woodland Survey Flora data to new schema
	GBW2	Code to transform Woodland Survey DBH data to new schema
	GBW3	Code to transform Woodland Survey plot information to new schema
	GBW4	Code to transform Woodland Survey soil information data to new schema
	GBW5	Code to incorporate Avery soil codes into Woodlands data
	GBW6	Code to transform Woodland Survey site data to new schema
	GBW7	Code to transform Woodland Survey site information data to new schema
	GBW8	Code to transform Woodland Survey data to new schema
Paper 6 (Key Habitats)	***KHS1***	Summary of procedure for creating mask/strata files for 'Key Habitat' Survey
	KHS2	Procedure for recreating the location of vegetation plots in the Key Habitat Survey
	KHS3	Examples of SAS code to update species data to address quality issues in plot data
	KHS4	Code to transpose Land Cover codes
	KHS5	Code to create Land Cover species attribute table
	KHS6	Code to transpose Boundary Codes
	KHS7	Transpose primary codes: Boundary species
	KHS8	Code to create perform updates on 'Key Habitat' survey Land Use codes
Papers 4,5,7 (Countryside Survey)	***CS1***	Script to update plot habitat allocation table with Broad Habitats allocated from digitized 1978 area data
	CS9	Countryside Survey 1978 Digitising Protocol

Dataset/Paper	Script ID	Description
Analyses		
Paper 1 (Pinewoods)	***SPW4***	Query to create table of the top 25 most frequently recorded ground flora species in pinewoods
	SPW5	Code to create charts showing tree diameter of Scots Pine (*Pinus sylvestris*) distribution in different woods
Papers 7 (Countryside Survey)	***CS2***	Code to create estimates of Broad Habitat area features in upland areas
	CS3	Code to create estimates of linear features in upland areas
	CS4	Code to summarise estimates of linear features by type and by country
	CS5	Code to summarise plots in upland areas
	CS6	Code to Summarise vegetation plots
	CS7	Summarise Countryside Survey vegetation plots by country
	CS8	Code to create estimates of grazed areas in the uplands

Table iii.1: List of documented scripts/methods

SPW1: Code to transpose diameter at breast height (DBH) codes

 # IMPORT DBH CODES TO SAS PROJECT

```
/*   START OF NODE: Import Data (Pinewoods1971.xls)   */

DATA WORK.IMPW_0003;
    LENGTH
        'Site No'n          8
        'Plot No'n          8
        Recorder          $ 10
        Date                8
        Quadrat             8
        'Tree/Saps/Shrub'n $ 17
        Species           $ 13
        Treeno              8
        DBH1                8
        DBH2                8
        DBH3                8
        DBH4                8
        DBH5                8
        DBH6                8
        DBH7                8
        DBH8                8
        DBH9                8
        DBH10               8
        DBH11               8
        DBH12               8
        DBH13               8
        DBH14               8
        DBH15               8
        DBH16               8
        DBH17               8
        DBH18               8
        DBH19               8
        DBH20               8
        DBH21               8
        DBH22               8
        DBH23               8
        DBH24               8
        DBH25               8
        DBH26               8
        DBH27               8
        DBH28               8
        DBH29               8
        DBH30               8
        DBH31               8
        DBH32               8
        DBH33               8
        DBH34               8
        DBH35               8
        DBH36               8
        DBH37               8
        DBH38               8
        DBH39               8
        DBH40               8
        DBH41               8
        DBH42               8
        DBH43               8
        DBH44               8
        DBH45               8
        DBH46               8
        DBH47               8
        DBH48               8
        DBH49               8
        DBH50               8
        DBH51               8
```

```
        DBH52              8 ;
    FORMAT
        Date              DATETIME18.0 ;
    INFORMAT
        Date              DATETIME18. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG9060\Pinewoods1971-
f62a2fe341404227a2d05e06cb96868d.txt'
        LRECL=250
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Site No'n        : BEST32.
        'Plot No'n        : BEST32.
        Recorder          :
        Date              : BEST32.
        Quadrat           : BEST32.
        'Tree/Saps/Shrub'n :
        Species           :
        Treeno            : BEST32.
        DBH1              : BEST32.
        DBH2              : BEST32.
        DBH3              : BEST32.
        DBH4              : BEST32.
        DBH5              : BEST32.
        DBH6              : BEST32.
        DBH7              : BEST32.
        DBH8              : BEST32.
        DBH9              : BEST32.
        DBH10             : BEST32.
        DBH11             : BEST32.
        DBH12             : BEST32.
        DBH13             : BEST32.
        DBH14             : BEST32.
        DBH15             : BEST32.
        DBH16             : BEST32.
        DBH17             : BEST32.
        DBH18             : BEST32.
        DBH19             : BEST32.
        DBH20             : BEST32.
        DBH21             : BEST32.
        DBH22             : BEST32.
        DBH23             : BEST32.
        DBH24             : BEST32.
        DBH25             : BEST32.
        DBH26             : BEST32.
        DBH27             : BEST32.
        DBH28             : BEST32.
        DBH29             : BEST32.
        DBH30             : BEST32.
        DBH31             : BEST32.
        DBH32             : BEST32.
        DBH33             : BEST32.
        DBH34             : BEST32.
        DBH35             : BEST32.
        DBH36             : BEST32.
        DBH37             : BEST32.
        DBH38             : BEST32.
        DBH39             : BEST32.
        DBH40             : BEST32.
        DBH41             : BEST32.
        DBH42             : BEST32.
        DBH43             : BEST32.
        DBH44             : BEST32.
        DBH45             : BEST32.
        DBH46             : BEST32.
```

```
        DBH47             : BEST32.
        DBH48             : BEST32.
        DBH49             : BEST32.
        DBH50             : BEST32.
        DBH51             : BEST32.
        DBH52             : BEST32. ;
RUN;
```

 # STACK ALL CODED DATA IN INDIVIDUAL COLUMNS INTO ONE COLUMN

```
/*   START OF NODE: Stack Columns   */
/*   Sort data set WORK.IMPW_0003--------------------------------------- */

PROC SORT
       DATA=WORK.IMPW_0003(KEEP=DBH1 DBH2 DBH3 DBH4 DBH5 DBH6 DBH7 DBH8 DBH9 DBH10
DBH11 DBH12 DBH13 DBH14 DBH15 DBH16 DBH17 DBH18 DBH19 DBH20 DBH21 DBH22 DBH23 DBH24
DBH25 DBH26 DBH27 DBH28 DBH29 DBH30 DBH31 DBH32 DBH33
         DBH34 DBH35 DBH36 DBH37 DBH38 DBH39 DBH40 DBH41 DBH42 DBH43 DBH44 DBH45
DBH46 DBH47 DBH48 DBH49 DBH50 DBH51 DBH52 "Site No"n "Plot No"n Treeno)
       OUT=WORK.TMP0TempTableInput
       ;
       BY "Site No"n "Plot No"n Treeno;
RUN;

PROC SQL;
       CREATE VIEW WORK.TMP1TempTableWork AS
       SELECT SRC.*, "StackedValues" AS _EG_IDCOL_
              FROM WORK.TMP0TempTableInput AS SRC;
QUIT;

PROC TRANSPOSE DATA = WORK.TMP1TempTableWork
       OUT=WORK.TRNSStackColumnsIMPW_0003(LABEL="Stacked WORK.IMPW_0003")
       NAME=ValueSource
       LABEL=ValueDescription
       ;
       BY "Site No"n "Plot No"n Treeno;
       ID _EG_IDCOL_;
       VAR DBH1 DBH2 DBH3 DBH4 DBH5 DBH6 DBH7 DBH8 DBH9 DBH10 DBH11 DBH12 DBH13
DBH14 DBH15 DBH16 DBH17 DBH18 DBH19 DBH20 DBH21 DBH22 DBH23 DBH24 DBH25 DBH26 DBH27
DBH28 DBH29 DBH30 DBH31 DBH32 DBH33 DBH34 DBH35 DBH36 DBH37 DBH38 DBH39 DBH40 DBH41
DBH42
         DBH43 DBH44 DBH45 DBH46 DBH47 DBH48 DBH49 DBH50 DBH51 DBH52;

RUN;

PROC DATASETS LIB=WORK NOLIST;
       MODIFY TRNSStackColumnsIMPW_0003;
       LABEL StackedValues = "The values of the columns being stacked.";
       LABEL ValueSource = "The name of the column from which the value came.";
       LABEL ValueDescription = "The label of the column from which the value
came.";
RUN;

/* ---- End of task code------------------------------ */
RUN; QUIT;
```

 # QUERY STACKED DATA TO FILTER OUT NULL VALUES. (EXPORT TABLE AND USE TEXT TO COLUMNS TO EXTRACT D = TREE DEAD DATA)

```
/*   START OF NODE: Query for Stacked WORK.IMPW_0003   */

PROC SQL;
   CREATE TABLE WORK.QUERY1099 AS
   SELECT TRNSSTACKCOLUMNSIMPW_0003.'Site No'n LABEL="Site No",
          TRNSSTACKCOLUMNSIMPW_0003.'Plot No'n LABEL="Plot No",
          TRNSSTACKCOLUMNSIMPW_0003.Treeno LABEL="Treeno",
          TRNSSTACKCOLUMNSIMPW_0003.StackedValues
      FROM WORK.TRNSSTACKCOLUMNSIMPW_0003 TRNSSTACKCOLUMNSIMPW_0003
```

```
      WHERE TRNSSTACKCOLUMNSIMPW_0003.StackedValues NOT IS NULL;
QUIT;
```

 # EXTRACT GROUPING COLUMNS FROM IMPORTED DBH DATA

```
/*   START OF NODE: Query for WORK.IMPW_0003   */

PROC SQL;
   CREATE TABLE WORK.Key AS
   SELECT DISTINCT IMPW_0003.'Site No'n LABEL="Site No",
          IMPW_0003.'Plot No'n LABEL="Plot No",
          IMPW_0003.Recorder LABEL="Recorder",
          IMPW_0003.Date LABEL="Date",
          IMPW_0003.Quadrat LABEL="Quadrat",
          IMPW_0003.'Tree/Saps/Shrub'n LABEL="Tree/Saps/Shrub",
          IMPW_0003.Species LABEL="Species",
          IMPW_0003.Treeno LABEL="Treeno"
      FROM WORK.IMPW_0003 IMPW_0003
      GROUP BY IMPW_0003.'Site No'n,
               IMPW_0003.'Plot No'n,
               IMPW_0003.Recorder,
               IMPW_0003.Date,
               IMPW_0003.Quadrat,
               IMPW_0003.'Tree/Saps/Shrub'n,
               IMPW_0003.Species,
               IMPW_0003.Treeno;
QUIT;
```

 # IMPORT TREE DEAD DATA INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Pinewoods1971.xls)   */

DATA WORK.IMPW_0004;
    LENGTH
        Site               8
        Plot               8
        TreeNO             8
        DBH                8
        Dead             $ 6 ;
    FORMAT
        Site             8.
        Plot             8.
        TreeNO           8.
        DBH              8.
        Dead             $CHAR6. ;
    INFORMAT
        Site             8.
        Plot             8.
        TreeNO           8.
        DBH              8.
        Dead             $CHAR6. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG9060\Pinewoods1971-
4c917c3639214f08acc2de2830e46503.txt'
        LRECL=17
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Site             : BEST32.
        Plot             : BEST32.
        TreeNO           : BEST32.
        DBH              : BEST32.
        Dead             : $CHAR6. ;
RUN;
```

 # JOIN SITE LIST TO TRANSPOSED DBH DATA TO CREATE OUTPUT

```
/*   START OF NODE: Query for WORK.IMPW_0004   */

PROC SQL;
   CREATE TABLE WORK.QUERY1531_0000 AS
   SELECT IMPW_0004.Site LABEL="Site",
          IMPW_0004.Plot LABEL="Plot",
          IMPW_0004.TreeNO LABEL="TreeNO",
          KEY.Recorder,
          KEY.Date,
          KEY.Quadrat,
          KEY.'Tree/Saps/Shrub'n,
          KEY.Species,
          IMPW_0004.DBH LABEL="DBH",
          IMPW_0004.Dead LABEL="Dead"
      FROM WORK.IMPW_0004 IMPW_0004, WORK.KEY KEY
      WHERE (IMPW_0004.TreeNO = KEY.Treeno AND IMPW_0004.Site = KEY.'Site No'n AND
IMPW_0004.Plot = KEY.'Plot No'n)
      ORDER BY IMPW_0004.TreeNO;
QUIT;
```

SPW2: Code to incorporate additional information into plot ground flora table

 # IMPORT FLORA DATA FROM EXCEL SHEET INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Pinewoods71_Ground_Flora.xlsx[Ground_Flora])   */

DATA WORK.Pinewoods71_Ground_Flora;
    LENGTH
        Table_id           8
        Data_id            8
        Site_no            8
        Plot_no            8
        Nest               8
        Code               8
        BRC                8
        Bryo             $ 1
        Code_PC            8
        Species          $ 41 ;
    FORMAT
        Table_id         BEST12.
        Data_id          BEST12.
        Site_no          BEST12.
        Plot_no          BEST12.
        Nest             BEST12.
        Code             BEST12.
        BRC              BEST12.
        Bryo             $CHAR1.
        Code_PC          BEST12.
        Species          $CHAR41. ;
    INFORMAT
        Table_id         BEST12.
        Data_id          BEST12.
        Site_no          BEST12.
        Plot_no          BEST12.
        Nest             BEST12.
        Code             BEST12.
        BRC              BEST12.
        Bryo             $CHAR1.
        Code_PC          BEST12.
        Species          $CHAR41. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13068\Pinewoods71_Ground_Flora-
6de872e300134eb9a82009fdf505ba39.txt'
```

```
        LRECL=74
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Table_id         : BEST32.
        Data_id          : BEST32.
        Site_no          : BEST32.
        Plot_no          : BEST32.
        Nest             : BEST32.
        Code             : BEST32.
        BRC              : BEST32.
        Bryo             : $CHAR1.
        Code_PC          : BEST32.
        Species          : $CHAR41. ;
RUN;
```

 # IMPORT SPECIES LIST FROM EXCEL SHEET

```
/*   START OF NODE: Import Data (Pinewoods71_Ground_Flora.xlsx[Sp_list])   */
DATA WORK.SP_LIST;
    LENGTH
        Table_id           8
        Code               8
        Description      $ 32
        Code_group       $ 28
        Code_group_description $ 10
        Data_sheet       $ 12 ;
    FORMAT
        Table_id         BEST12.
        Code             BEST12.
        Description      $CHAR32.
        Code_group       $CHAR28.
        Code_group_description $CHAR10.
        Data_sheet       $CHAR12. ;
    INFORMAT
        Table_id         BEST12.
        Code             BEST12.
        Description      $CHAR32.
        Code_group       $CHAR28.
        Code_group_description $CHAR10.
        Data_sheet       $CHAR12. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13068\Pinewoods71_Ground_Flora-
a22c7f6c5d524286bc953fa9091ae9b7.txt'
        LRECL=85
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Table_id         : BEST32.
        Code             : BEST32.
        Description      : $CHAR32.
        Code_group       : $CHAR28.
        Code_group_description : $CHAR10.
        Data_sheet       : $CHAR12. ;
RUN;
```

 # JOIN RECORDED SPECIES TO DESCRIPTIONS

```
/*   START OF NODE: Query Builder   */

PROC SQL;
```

```
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUND_FLO(label="QUERY_FOR_PINEWOODS71_GROUND_FLORA")
AS
   SELECT DISTINCT t1.Table_id,
          t1.Data_id,
          t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.Code,
          t1.BRC,
          t1.Bryo,
          t1.Code_PC,
          t1.Species,
          t2.Table_id AS Table_id1,
          t2.Code AS Code1,
          t2.Description,
          t2.Code_group,
          t2.Code_group_description,
          t2.Data_sheet
      FROM WORK.PINEWOODS71_GROUND_FLORA t1
           LEFT JOIN WORK.SP_LIST t2 ON (t1.Code = t2.Code);
QUIT;
```

 # UPDATE SPECIES DESCRIPTIONS IN LIST

```
/*   START OF NODE: Query Builder1   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_0000(label="QUERY_FOR_PINEWOODS71_GROUND_FLO") AS
   SELECT t1.Table_id,
          t1.Data_id,
          t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.Code,
          t1.BRC,
          t1.Bryo,
          t1.Code_PC,
          t1.Species,
          t1.Table_id1,
          t1.Code1,
          t1.Description,
          t1.Code_group,
          t1.Code_group_description,
          t1.Data_sheet,
          /* SP_ALL */
            (CASE
               WHEN t1.Description is null
               THEN t1.Species
               ELSE t1.Description
            END) AS SP_ALL
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUND_FLO t1;
QUIT;
```

 # UPDATE BRYOPHYTE SPECIES WITH NO DESCRIPTION

```
/*   START OF NODE: Query Builder2   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_0003(label="QUERY_FOR_PINEWOODS71_GROUN") AS
   SELECT DISTINCT t1.Table_id,
          t1.Data_id,
          t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.Code,
```

```
          t1.BRC,
          t1.Bryo,
          t1.Code_PC,
          t1.Species,
          t1.Table_id1,
          t1.Code1,
          t1.Description,
          t1.Code_group,
          t1.Code_group_description,
          t1.Data_sheet,
          t1.SP_ALL,
          /* SP_ALL_2 */
            (CASE
               WHEN t1.SP_ALL is null
               THEN "Bryophyte sp."
               ELSE t1.SP_ALL
            END) AS SP_ALL_2
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_0000 t1;
QUIT;
```

 # IDENTIFY SPECIES WITH MISSING DESCRIPTIONS IN LIST

```
/*   START OF NODE: Query Builder3   */
PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_0004(label="QUERY_FOR_PINEWOODS71_GROUN") AS
   SELECT DISTINCT t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.Code,
          t1.Code_PC,
          t1.Description,
          t1.SP_ALL_2,
          t2.BRC_NAMES,
          t2.BRC_NUMBER,
          t2.COMMON_NAME
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_0003 t1
           LEFT JOIN CS_VEG.LUS_SP_LIB_AND_TRAITS t2 ON (t1.SP_ALL_2 =
t2.BRC_NAMES);
QUIT;


/*   START OF NODE: Query Builder4   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_0005(label="QUERY_FOR_PINEWOODS71_GROUN") AS
   SELECT DISTINCT t1.SP_ALL_2,
          t1.BRC_NUMBER
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_0004 t1
      WHERE t1.BRC_NAMES IS MISSING;
QUIT;
```

 # EXPORT LIST OF MISSING SPECIES AND CROSS CHECK WITH FIELD SHEETS. REIMPORT COMPLETED LIST INTO SAS PROJECT

```
/* START OF NODE: Import Data (Pinewoods71_Ground_Flora.xlsx[Missing _Codes])   */

DATA WORK.Pinewoods71_Ground_Flora1;
    LENGTH
        F1               $ 41
        F2                 8 ;
    FORMAT
        F1               $CHAR41.
        F2               BEST12. ;
    INFORMAT
        F1               $CHAR41.
        F2               BEST12. ;
```

```
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13068\Pinewoods71_Ground_Flora-
1c28ab49ae024892bb1bcd0e20fd2ef2.txt'
        LRECL=49
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        F1               : $CHAR41.
        F2               : BEST32. ;
RUN;
```

 # JOIN DATA TO UPDATED SPECIES LOOKUP

```
/*   START OF NODE: Query Builder6   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_0006(label="QUERY_FOR_PINEWOODS71_GROUN") AS
   SELECT t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.Code,
          t1.Code_PC,
          t1.Description,
          t1.SP_ALL_2,
          t1.BRC_NAMES,
          t1.BRC_NUMBER,
          t1.COMMON_NAME,
          t2.F1,
          t2.F2,
          /* BRC_ALL */
            (CASE
               WHEN t1.BRC_NUMBER is null
               THEN t2.F2
               ELSE t1.BRC_NUMBER
            END) AS BRC_ALL
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_0004 t1
           LEFT JOIN WORK.PINEWOODS71_GROUND_FLORA1 t2 ON (t1.SP_ALL_2 = t2.F1);
QUIT;

/*   START OF NODE: Query Builder7   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_0007(label="QUERY_FOR_PINEWOODS71_GROUN") AS
   SELECT DISTINCT t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.Code,
          t1.Code_PC,
          t1.Description,
          t1.SP_ALL_2,
          t1.BRC_NAMES,
          t1.BRC_NUMBER,
          t1.COMMON_NAME,
          t1.F1,
          t1.F2,
          t1.BRC_ALL,
          t2.BRC_NAMES AS BRC_NAMES1,
          t2.BRC_NUMBER AS BRC_NUMBER1,
          t2.COMMON_NAME AS COMMON_NAME1
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_0006 t1
           LEFT JOIN CS_VEG.LUS_SP_LIB_AND_TRAITS t2 ON (t1.BRC_ALL =
t2.BRC_NUMBER);
QUIT;
```

```
/*   START OF NODE: Query Builder8   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_0008(label="QUERY_FOR_PINEWOODS71_GROUN") AS
   SELECT DISTINCT t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.Code AS ORIG_CODE,
          t1.Code_PC AS COVER,
          t1.SP_ALL_2 AS ORIG_SP_NAME,
          t1.BRC_NUMBER1 AS BRC_NUMBER,
          t1.BRC_NAMES1 AS BRC_NAMES,
          t1.COMMON_NAME1 AS COMMON_NAME
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_0007 t1;
QUIT;

/*   START OF NODE: Query Builder9   */
PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_0009(label="QUERY_FOR_PINEWOODS71_GROUN") AS
   SELECT t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.ORIG_CODE,
          t1.COVER,
          t1.ORIG_SP_NAME,
          t1.BRC_NUMBER,
          t1.BRC_NAMES,
          t1.COMMON_NAME
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_0008 t1
      WHERE t1.BRC_NAMES IS MISSING;
QUIT;

/*   START OF NODE: Query Builder10   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_PINEWOODS71_GROUN_000B(label="QUERY_FOR_PINEWOODS71_GROUN") AS
   SELECT DISTINCT t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.COVER,
          t1.ORIG_CODE,
          t1.ORIG_SP_NAME,
          t1.BRC_NUMBER,
          /* BRC_NAME_FINAL */
            (CASE
               WHEN t1.BRC_NAMES is null
               THEN t1.ORIG_SP_NAME
               ELSE t1.BRC_NAMES
            END) AS BRC_NAME_FINAL,
          t1.COMMON_NAME
      FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_0008 t1;
QUIT;
```

 # IMPORT BARE GROUND/LITTER DATA

```
/*   START OF NODE: Import Data (Plot_bryo_etc.xlsx[Sheet1])   */

DATA WORK.Plot_bryo_etc;
    LENGTH
        Site_no            8
        Plot_no            8
        Plot_Litter        8
        Plot_wood          8
        Plot_rock          8
        Plot_bare          8
```

```
        Plot_water         8
        Bryophytes         8 ;
    FORMAT
        Site_no          BEST12.
        Plot_no          BEST12.
        Plot_Litter      BEST12.
        Plot_wood        BEST12.
        Plot_rock        BEST12.
        Plot_bare        BEST12.
        Plot_water       BEST12.
        Bryophytes       BEST12. ;
    INFORMAT
        Site_no          BEST12.
        Plot_no          BEST12.
        Plot_Litter      BEST12.
        Plot_wood        BEST12.
        Plot_rock        BEST12.
        Plot_bare        BEST12.
        Plot_water       BEST12.
        Bryophytes       BEST12. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13068\Plot_bryo_etc-
f5f0582186b34f86909ee6e4ce0a4a9e.txt'
        LRECL=27
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Site_no          : BEST32.
        Plot_no          : BEST32.
        Plot_Litter      : BEST32.
        Plot_wood        : BEST32.
        Plot_rock        : BEST32.
        Plot_bare        : BEST32.
        Plot_water       : BEST32.
        Bryophytes       : BEST32. ;
RUN;
```

 # RENAME COLUMNS

```
/*   START OF NODE: Query Builder16   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_PLOT_BRYO_ETC AS
   SELECT DISTINCT t1.Site_no,
          t1.Plot_no,
          t1.Plot_Litter AS TOTAL_LITTER,
          t1.Plot_wood AS TOTAL_WOOD,
          t1.Plot_rock AS TOTAL_ROCK,
          t1.Plot_bare AS TOTAL_BARE,
          t1.Plot_water AS TOTAL_WATER,
          t1.Bryophytes AS TOTAL_BRYO
      FROM WORK.PLOT_BRYO_ETC t1;
QUIT;
```

 # TRANSPOSE COLUMNS INTO ONE COLUMN

```
/*   START OF NODE: Transpose   */
/*    Sort data set Local:WORK.QUERY_FOR_PLOT_BRYO_ETC----- */

PROC SQL;
       CREATE VIEW WORK.SORTTempTableSorted AS
              SELECT T.TOTAL_LITTER, T.TOTAL_WOOD, T.TOTAL_ROCK, T.TOTAL_BARE,
T.TOTAL_WATER, T.TOTAL_BRYO, T.Site_no, T.Plot_no
       FROM WORK.QUERY_FOR_PLOT_BRYO_ETC as T
;
QUIT;
```

```
PROC TRANSPOSE DATA=WORK.SORTTempTableSorted
       OUT=WORK.TRNSTransposedQUERY_FOR_PLOT_BRY(LABEL="Transposed
WORK.QUERY_FOR_PLOT_BRYO_ETC")
       PREFIX=Column
       NAME=Source
       LABEL=Label
;
       BY Site_no Plot_no;
       VAR TOTAL_LITTER TOTAL_WOOD TOTAL_ROCK TOTAL_BARE TOTAL_WATER TOTAL_BRYO;

/* ---  End of task code--------------------------------------------- */

RUN; QUIT;
```

 # RENAME COLUMNS AND FILTER OUT NULLS

```
/*   START OF NODE: Query Builder17   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_TRNSTRANSPOSEDQUERY_FO(label="QUERY_FOR_TRNSTRANSPOSEDQUERY_FOR_PLOT
_BRY") AS
   SELECT DISTINCT t1.Site_no,
          t1.Plot_no,
          t1.Source AS BRC_NAME_FINAL,
          t1.Column1 AS COVER,
          /* NEST */
            (5) AS NEST
      FROM WORK.TRNSTRANSPOSEDQUERY_FOR_PLOT_BRY t1
      WHERE t1.Column1 NOT IS MISSING;
QUIT;
```

 # APPEND BARE GROUND/LITTER ETC. TO SPECIES TABLE

```
/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.Append_Table AS
SELECT * FROM WORK.QUERY_FOR_PINEWOODS71_GROUN_000B
 OUTER UNION CORR
SELECT * FROM WORK.QUERY_FOR_TRNSTRANSPOSEDQUERY_FO
;
Quit;
```

 # CREATE SPECIES TABLE

```
/*   START OF NODE: Query Builder18   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_APPEND_TABLE AS
   SELECT t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.COVER,
          t1.ORIG_CODE,
          t1.ORIG_SP_NAME,
          t1.BRC_NUMBER,
          t1.BRC_NAME_FINAL,
          t1.COMMON_NAME
      FROM WORK.APPEND_TABLE t1
      ORDER BY t1.Site_no,
               t1.Plot_no,
               t1.Nest;
QUIT;
```

 # APPEND ADDITIONAL SPECIES INFORMATION FROM CEH DATABASE

```
/*   START OF NODE: Query Builder23   */
```

```
PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_APPEND_TABLE_0004(label="QUERY_FOR_APPEND_TABLE") AS
   SELECT t1.Site_no,
          t1.Plot_no,
          t1.Nest,
          t1.COVER,
          t1.ORIG_CODE,
          t1.ORIG_SP_NAME,
          t1.BRC_NUMBER,
          t1.BRC_NAME_FINAL,
          t2.COMMON_NAME,
          t2.GROWTH_FORM
      FROM QUERY_FOR_APPEND_TABLE t1
           LEFT JOIN CS_VEG.LUS_SP_LIB_AND_TRAITS t2 ON (t1.BRC_NUMBER =
t2.BRC_NUMBER)
      ORDER BY t1.Site_no,
               t1.Plot_no,
               t1.Nest;
QUIT;
```

SPW3 Code to transpose plot description codes (soil)

 # IMPORT CODES TO SAS PROJECT

```
/*   START OF NODE: Import Data (soil_codes.csv)   */

DATA WORK.IMPW_0000;
    LENGTH
        'Site No'n         8
        'Plot No'n         8
        Code1              8
        Code2              8
        Code3              8
        Code4              8
        Code5              8
        Code6              8
        Code7              8
        Code8              8
        Code9              8
        Code10             8
        Code11             8
        Code12             8
        Code13             8
        Code14             8
        Code15             8
        Code16             8
        Code17             8
        Code18             8
        Code19             8
        Code20             8
        Code21             8
        Code22             8 ;
    INFORMAT
        'Site No'n       COMMA32.
        'Plot No'n       COMMA32.
        Code1            COMMA32.
        Code2            COMMA32.
        Code3            COMMA32.
        Code4            COMMA32.
        Code5            COMMA32.
        Code6            COMMA32.
        Code7            COMMA32.
```

```
        Code8            COMMA32.
        Code9            COMMA32.
        Code10           COMMA32.
        Code11           COMMA32.
        Code12           COMMA32.
        Code13           COMMA32.
        Code14           COMMA32.
        Code15           COMMA32.
        Code16           COMMA32.
        Code17           COMMA32.
        Code18           COMMA32.
        Code19           COMMA32.
        Code20           COMMA32.
        Code21           COMMA32.
        Code22           COMMA32. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG9060\soil_codes-
f3674f343c2c45c3bfdcbee4f5e48fcf.txt'
        LRECL=67
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Site No'n       : ?? COMMA32.
        'Plot No'n       : ?? COMMA32.
        Code1            : ?? COMMA32.
        Code2            : ?? COMMA32.
        Code3            : ?? COMMA32.
        Code4            : ?? COMMA32.
        Code5            : ?? COMMA32.
        Code6            : ?? COMMA32.
        Code7            : ?? COMMA32.
        Code8            : ?? COMMA32.
        Code9            : ?? COMMA32.
        Code10           : ?? COMMA32.
        Code11           : ?? COMMA32.
        Code12           : ?? COMMA32.
        Code13           : ?? COMMA32.
        Code14           : ?? COMMA32.
        Code15           : ?? COMMA32.
        Code16           : ?? COMMA32.
        Code17           : ?? COMMA32.
        Code18           : ?? COMMA32.
        Code19           : ?? COMMA32.
        Code20           : ?? COMMA32.
        Code21           : ?? COMMA32.
        Code22           : ?? COMMA32. ;
RUN;
```

 # STACK ALL CODED DATA IN INDIVIDUAL COLUMNS INTO ONE COLUMN

```
/*   START OF NODE: Stack Columns   */
/* ---Sort data set WORK.IMPW_0000----------------------------------- */

PROC SORT
       DATA=WORK.IMPW_0000(KEEP=Code1 Code2 Code3 Code4 Code5 Code6 Code7 Code8
Code9 Code10 Code11 Code12 Code13 Code14 Code15 Code16 Code17 Code18 Code19 Code20
Code21 Code22 "Site No"n "Plot No"n)
       OUT=WORK.TMP0TempTableInput
       ;
       BY "Site No"n "Plot No"n;
RUN;

PROC SQL;
       CREATE VIEW WORK.TMP1TempTableWork AS
       SELECT SRC.*, "StackedValues" AS _EG_IDCOL_
              FROM WORK.TMP0TempTableInput AS SRC;
```

```
QUIT;

PROC TRANSPOSE DATA = WORK.TMP1TempTableWork
       OUT=WORK.TRNSStackColumnsIMPW_0000(LABEL="Stacked WORK.IMPW_0000")
       NAME=ValueSource
       LABEL=ValueDescription
       ;
       BY "Site No"n "Plot No"n;
       ID _EG_IDCOL_;
       VAR Code1 Code2 Code3 Code4 Code5 Code6 Code7 Code8 Code9 Code10 Code11
Code12 Code13 Code14 Code15 Code16 Code17 Code18 Code19 Code20 Code21 Code22;

RUN;

PROC DATASETS LIB=WORK NOLIST;
       MODIFY TRNSStackColumnsIMPW_0000;
       LABEL StackedValues = "The values of the columns being stacked.";
       LABEL ValueSource = "The name of the column from which the value came.";
       LABEL ValueDescription = "The label of the column from which the value
came.";
RUN;

/* - End of task code--------------------------------- */
```

 # QUERY STACKED DATA TO FILTER OUT NULL VALUES

```
/*   START OF NODE: Query for Stacked WORK.IMPW_0000   */

PROC SQL;
   CREATE TABLE WORK.QUERY8408 AS
   SELECT TRNSSTACKCOLUMNSIMPW_0000.'Site No'n LABEL="Site No",
          TRNSSTACKCOLUMNSIMPW_0000.'Plot No'n LABEL="Plot No",
          TRNSSTACKCOLUMNSIMPW_0000.StackedValues
      FROM WORK.TRNSSTACKCOLUMNSIMPW_0000 TRNSSTACKCOLUMNSIMPW_0000
      WHERE TRNSSTACKCOLUMNSIMPW_0000.StackedValues NOT IS NULL;
QUIT;
```

GBW1: Code to transform Woodland Survey Flora data to new schema

(Final file: Woodlands_Survey_Flora_Data_1971_2001.csv)

 # IMPORT FLORA DATA INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.IMPW_0001;
    LENGTH
        SITE               8
        PLOT               8
        BRC_number         8
        NEST               8
        COV                8
        YR                 8
        Bryo             $ 6
        Amalgams           8
        Yr_2               8
        'BRC names'n     $ 27
        Woody            $ 7 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
caec1df0500e43c083e257d910a2e4bb.txt'
        LRECL=83
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
```

```
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        BRC_number       : BEST32.
        NEST             : BEST32.
        COV              : BEST32.
        YR               : BEST32.
        Bryo             :
        Amalgams         : BEST32.
        Yr_2             : BEST32.
        'BRC names'n     :
        Woody            :  ;
RUN;
```

 # IMPORT SPECIES LIST INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.SP_LIB;
    LENGTH
        'BRC names'n     $ 73
        'BRC number'n      8
        Woody            $ 7 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
1abec44315ab4f8ab08443758adc0db8.txt'
        LRECL=79
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'BRC names'n     :
        'BRC number'n    : BEST32.
        Woody            :  ;
RUN;
```

 # IMPORT DATA INTO SAS PROJECT (BARE GROUND/LITTER ETC 1971)

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.IMPW_0003;
    LENGTH
        SITE               8
        PLOT               8
        Code               8
        Cover              8
        NEST               8
        YR                 8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
41f4a187c8c948f58d5a66e985582720.txt'
        LRECL=28
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        Code             : BEST32.
        Cover            : BEST32.
        NEST             : BEST32.
        YR               : BEST32. ;
RUN;
```

 # RENAME COLUMNS AND JOIN RECORDED CODED DATA TO SPECIES LIST

```
/*   START OF NODE: Query for WORK.IMPW_0003   */
PROC SQL;
   CREATE TABLE WORK.QUERY1706 AS
   SELECT IMPW_0003.SITE LABEL="SITE",
          IMPW_0003.PLOT LABEL="PLOT",
          IMPW_0003.Code LABEL="Code",
          IMPW_0003.Cover LABEL="Cover",
          IMPW_0003.NEST LABEL="NEST",
          IMPW_0003.YR LABEL="YR",
          SP_LIB.'BRC names'n LABEL="BRC names"
      FROM WORK.IMPW_0003 IMPW_0003
           INNER JOIN WORK.SP_LIB SP_LIB ON (IMPW_0003.Code = SP_LIB.'BRC
number'n);
QUIT;
```

 # IMPORT DATA INTO SAS PROJECT (BARE GROUND/LITTER ETC 2003)

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.IMPW_0004;
    LENGTH
        SITE               8
        PLOT               8
        Cover              8
        code               8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
92839a09abde4b1c9b2c28336600e34e.txt'
        LRECL=21
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        Cover            : BEST32.
        code             : BEST32. ;
RUN;
```

 # RENAME COLUMNS AND JOIN RECORDED CODED DATA TO SPECIES LIST (BARE GROUND/LITTER ETC)

```
/*   START OF NODE: Query for WORK.IMPW_0004   */

PROC SQL;
   CREATE TABLE WORK.QUERY136 AS
   SELECT IMPW_0004.SITE LABEL="SITE",
          IMPW_0004.PLOT LABEL="PLOT",
          IMPW_0004.Cover LABEL="Cover",
          IMPW_0004.code LABEL="code",
          SP_LIB.'BRC names'n LABEL="BRC names",
          /* YR */
            (1971) AS YR
      FROM WORK.IMPW_0004 IMPW_0004
           INNER JOIN WORK.SP_LIB SP_LIB ON (IMPW_0004.code = SP_LIB.'BRC
number'n);
QUIT;
```

 # APPEND RECORDED DATA TO BARE GROUND/LITTER ETC DATA

```
/*   START OF NODE: Append Table   */

%_eg_conditional_dropds(WORK.APPEND_TABLE_0002);
PROC SQL;
CREATE TABLE WORK.APPEND_TABLE_0002 AS
SELECT * FROM WORK.QUERY1706
 OUTER UNION CORR
```

```
SELECT * FROM WORK.QUERY136
;
Quit;

/*   START OF NODE: Query for APPEND_TABLE_0002   */

PROC SQL;
   CREATE TABLE WORK.Query_for_APPEND_TABLE_0002 AS
   SELECT APPEND_TABLE_0002.SITE,
          APPEND_TABLE_0002.PLOT,
          APPEND_TABLE_0002.Code AS BRC_number,
          APPEND_TABLE_0002.Cover AS COV,
          APPEND_TABLE_0002.NEST,
          APPEND_TABLE_0002.YR,
          APPEND_TABLE_0002.'BRC names'n
      FROM WORK.APPEND_TABLE_0002 APPEND_TABLE_0002;
QUIT;

/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.APPEND_TABLE_0004 AS
SELECT * FROM WORK.IMPW_0001
 OUTER UNION CORR
SELECT * FROM WORK.QUERY_FOR_APPEND_TABLE_0002
;
Quit;
```

 # RENAME COLUMNS AND PRODUCE FINAL OUTPUT TABLE

```
/*   START OF NODE: Query for APPEND_TABLE_0004   */

PROC SQL;
   CREATE TABLE WORK.Query_for_APPEND_TABLE_0004 AS
   SELECT APPEND_TABLE_0004.SITE,
          APPEND_TABLE_0004.PLOT,
          APPEND_TABLE_0004.BRC_number LABEL="BRC_number",
          APPEND_TABLE_0004.NEST,
          APPEND_TABLE_0004.COV LABEL="COV",
          APPEND_TABLE_0004.YR,
          APPEND_TABLE_0004.Bryo LABEL="Bryo",
          APPEND_TABLE_0004.Amalgams LABEL="Amalgams",
          APPEND_TABLE_0004.Yr_2 LABEL="Yr_2",
          APPEND_TABLE_0004.'BRC names'n,
          APPEND_TABLE_0004.Woody LABEL="Woody",
          SP_LIB.'BRC names'n LABEL="BRC names" AS 'BRC namesaggs (BRC names)'n
      FROM WORK.APPEND_TABLE_0004 APPEND_TABLE_0004
           INNER JOIN WORK.SP_LIB SP_LIB ON (APPEND_TABLE_0004.Amalgams =
SP_LIB.'BRC number'n)
      WHERE APPEND_TABLE_0004.BRC_number <> APPEND_TABLE_0004.Amalgams;
QUIT;

/*   START OF NODE: Query for Query_for_APPEND_TABLE_0004   */

PROC SQL;
   CREATE TABLE WORK.QUERY6115 AS
   SELECT DISTINCT QUERY_FOR_APPEND_TABLE_0004.BRC_number,
          QUERY_FOR_APPEND_TABLE_0004.Amalgams,
          QUERY_FOR_APPEND_TABLE_0004.'BRC names'n,
          QUERY_FOR_APPEND_TABLE_0004.'BRC namesaggs (BRC names)'n
      FROM WORK.QUERY_FOR_APPEND_TABLE_0004 QUERY_FOR_APPEND_TABLE_0004
      GROUP BY QUERY_FOR_APPEND_TABLE_0004.BRC_number,
               QUERY_FOR_APPEND_TABLE_0004.Amalgams,
               QUERY_FOR_APPEND_TABLE_0004.'BRC names'n,
               QUERY_FOR_APPEND_TABLE_0004.'BRC namesaggs (BRC names)'n;
QUIT;
```

GBW2: Code to transform Woodland Survey DBH data to new schema

(Final table: Woodlands_Survey_Tree_Diameter_Data_1971_2001.csv)

 # IMPORT DBH DATA (DEAD TREES) INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.DBHDEAD;
    LENGTH
        SITE               8
        PLOT               8
        BRC_code           8
        DBHclass           8
        Final_count        8
        Yr                 8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
0ceaa5ba342247e793cc635f8bdbdb11.txt'
        LRECL=21
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        BRC_code         : BEST32.
        DBHclass         : BEST32.
        Final_count      : BEST32.
        Yr               : BEST32. ;
RUN;
```

 # RENAME COLUMNS

```
/*   START OF NODE: Query for WORK.DBHDEAD   */

PROC SQL;
   CREATE TABLE WORK.QUERY393 AS
   SELECT DBHDEAD.SITE LABEL="SITE",
          DBHDEAD.PLOT LABEL="PLOT",
          DBHDEAD.BRC_code LABEL="BRC_code" AS BRC_number,
          DBHDEAD.DBHclass LABEL="DBHclass",
          DBHDEAD.Final_count LABEL="Final_count",
          DBHDEAD.Yr LABEL="Yr",
          /* Status */
            ("Dead") AS Status
      FROM WORK.DBHDEAD DBHDEAD;
QUIT;
```

 # IMPORT DBH DATA (LIVE TREES) INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.DBHLIVE;
    LENGTH
        SITE               8
        PLOT               8
        Amalgams           8
        DBH_class          8
        Count              8
        BRC_number         8
```

```
        Yr                 8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
36c741219f834715b1426e1bea54a70d.txt'
        LRECL=33
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        Amalgams         : BEST32.
        DBH_class        : BEST32.
        Count            : BEST32.
        BRC_number       : BEST32.
        Yr               : BEST32. ;
RUN;
```

 # RENAME COLUMNS

```
/*   START OF NODE: Query for WORK.DBHLIVE   */

PROC SQL;
   CREATE TABLE WORK.QUERY3059 AS
   SELECT DBHLIVE.SITE LABEL="SITE",
          DBHLIVE.PLOT LABEL="PLOT",
          DBHLIVE.Amalgams LABEL="Amalgams",
          DBHLIVE.DBH_class LABEL="DBH_class" AS 'DBHclass 'n,
          DBHLIVE.Count LABEL="Count",
          DBHLIVE.BRC_number LABEL="BRC_number",
          DBHLIVE.Yr LABEL="Yr",
          /* Status */
            ("Live") AS Status
      FROM WORK.DBHLIVE DBHLIVE;
QUIT;
```

 # APPEND LIVE AND DEAD DBH DATA

```
/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.APPEND_TABLE_0007 AS
SELECT * FROM WORK.QUERY3059
 OUTER UNION CORR
SELECT * FROM WORK.QUERY393
;
Quit;
```

 # ADD LATIN SPECIES NAMES TO DATA SET

```
/*   START OF NODE: Query for APPEND_TABLE_0007   */

PROC SQL;
   CREATE TABLE WORK.Query_for_APPEND_TABLE_0007 AS
   SELECT APPEND_TABLE_0007.SITE,
          APPEND_TABLE_0007.PLOT,
          APPEND_TABLE_0007.Amalgams,
          APPEND_TABLE_0007.DBHclass,
          APPEND_TABLE_0007.Count,
          APPEND_TABLE_0007.BRC_number,
          APPEND_TABLE_0007.Yr,
          APPEND_TABLE_0007.Status,
```

```
          APPEND_TABLE_0007.Final_count,
          SP_LIB.'BRC names'n LABEL="BRC names",
          SP_LIB.Woody LABEL="Woody",
          SP_LIB1.'BRC names'n LABEL="BRC names" AS 'BRC names1'n,
          SP_LIB1.Woody LABEL="Woody" AS Woody1
      FROM WORK.APPEND_TABLE_0007 APPEND_TABLE_0007, WORK.SP_LIB SP_LIB,
WORK.SP_LIB SP_LIB1
      WHERE (APPEND_TABLE_0007.Amalgams = SP_LIB.'BRC number'n AND
APPEND_TABLE_0007.BRC_number = SP_LIB1.'BRC number'n);
QUIT;
```

GBW3: Code to transform Woodland Survey plot information to new schema

(Part of Final table: Woodlands_Survey_Site_Information_1971_2001.csv)

 # IMPORT PLOT DATA 2003 INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.PLOTS03;
    LENGTH
        SITE               8
        PLOT               8
        Notes            $ 30
        PD_code            8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
6d2e3cb9f6ea41748946c4c3f4d147d0.txt'
        LRECL=251
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        Notes            :
        PD_code          : BEST32. ;
RUN;
```

 # IMPORT PLOT DATA 1971 INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.PLOTS71;
    LENGTH
        SITE               8
        PLOT               8
        Notes            $ 50
        PD_code            8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
b01a1405b3454f37bcb37d64ab006546.txt'
        LRECL=52
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        Notes            :
        PD_code          : BEST32. ;
RUN;
```

 # IMPORT PLOT CODE LIST INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.PLOTLIB;
    LENGTH
        'Plot code'n       8
        Description      $ 22 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
c04f31b205cc452ab2b8376b8d2df7d3.txt'
        LRECL=22
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Plot code'n     : BEST32.
        Description      :  ;
RUN;
```

 # IMPORT PLOT SLOPE DATA 1971 INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.SLOPE71;
    LENGTH
        SITE               8
        PLOT               8
        PD_code            8
        Slope_aspect     $ 14 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
9f8afc59954c4a98aec7841eef0acd79.txt'
        LRECL=12
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        PD_code          : BEST32.
        Slope_aspect     :  ;
RUN;
```

 # IMPORT PLOT SLOPE DATA 2003 INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.SLOPE03;
    LENGTH
        SITE               8
        PLOT               8
        PD_code            8
        Slope_aspect     $ 14 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
9f8afc59954c4a98aec7841eef0acd80.txt'
        LRECL=12
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        PD_code          : BEST32.
        Slope_aspect     :  ;
```

```
RUN;
```

 # SELECT SLOPE (CODE 5) 1971 INFORMATION FROM TABLE

```
/*   START OF NODE: Query for WORK.SLOPE71   */

PROC SQL;
   CREATE TABLE WORK.slopecode571 AS
   SELECT SLOPE71.SITE LABEL="SITE",
          SLOPE71.PLOT LABEL="PLOT",
          SLOPE71.PD_code LABEL="PD_code",
          SLOPE71.Slope_aspect LABEL="Slope_aspect",
          /* Year */
            ("1971") AS Year
      FROM WORK.SLOPE71 SLOPE71
      WHERE SLOPE71.PD_code = 5;
QUIT;
```

 # SELECT ASPECT (CODE 6) 1971 INFORMATION FROM TABLE

```
/*   START OF NODE: Query1 for WORK.SLOPE71   */

PROC SQL;
   CREATE TABLE WORK.aspect71 AS
   SELECT SLOPE71.SITE LABEL="SITE",
          SLOPE71.PLOT LABEL="PLOT",
          SLOPE71.PD_code LABEL="PD_code",
          SLOPE71.Slope_aspect LABEL="Slope_aspect",
          /* Year */
            ("1971") AS Year
      FROM WORK.SLOPE71 SLOPE71
      WHERE SLOPE71.PD_code = 6;
QUIT;
```

 # SELECT SLOPE (CODE 5) 2003 INFORMATION FROM TABLE

```
PROC SQL;
   CREATE TABLE WORK.slopecode503 AS
   SELECT SLOPE03.SITE LABEL="SITE",
          SLOPE03.Plot LABEL="Plot",
          SLOPE03.Slope_aspect LABEL="Slope_aspect",
          SLOPE03.PD_code LABEL="PD_code",
          /* Year */
            ("2003") AS Year
      FROM WORK.SLOPE03 SLOPE03
      WHERE SLOPE03.PD_code = 5;
QUIT;
```

 # SELECT ASPECT (CODE 6) 2003 INFORMATION FROM TABLE

```
PROC SQL;
   CREATE TABLE WORK.aspect03 AS
   SELECT SLOPE03.SITE LABEL="SITE",
          SLOPE03.Plot LABEL="Plot",
          SLOPE03.Slope_aspect LABEL="Slope_aspect",
          SLOPE03.PD_code LABEL="PD_code",
          /* Year */
            ("2003") AS Year
      FROM WORK.SLOPE03 SLOPE03
      WHERE SLOPE03.PD_code = 6;
QUIT;
```

 # EXTRACT DATA NOT CODED 5 OR 6 (SLOPE/ASPECT) FROM TABLE TO CHECK VALUES IN WRONG COLUMN

```
/*   START OF NODE: Query2 for WORK.SLOPE71   */

PROC SQL;
   CREATE TABLE WORK.QUERY7867 AS
```

```
   SELECT SLOPE71.SITE LABEL="SITE",
          SLOPE71.PLOT LABEL="PLOT",
          SLOPE71.PD_code LABEL="PD_code",
          SLOPE71.Slope_aspect LABEL="Slope_aspect"
      FROM WORK.SLOPE71 SLOPE71
      WHERE SLOPE71.PD_code NOT IN
           (
           5,
           6
           );
QUIT;
```

 # UPDATE SLOPE/ASPECT VALUES ENTERED IN THE WRONG COLUMN

```
/*   START OF NODE: Query for QUERY7867   */

PROC SQL;
   CREATE TABLE WORK.slopeextra71 AS
   SELECT QUERY7867.SITE,
          QUERY7867.PLOT,
          QUERY7867.PD_code AS Slope_aspect,
          QUERY7867.Slope_aspect AS PD_Code
      FROM WORK.QUERY7867 QUERY7867
      WHERE QUERY7867.Slope_aspect = '5';
QUIT;
```

 # CONVERT SLOPE/ASPECT VALUES FROM TEXT TO NUMERIC

```
/*   START OF NODE: Query for slopeextra71   */

PROC SQL;
   CREATE TABLE WORK.Query_for_slopeextra71 AS
   SELECT SLOPEEXTRA71.SITE,
          SLOPEEXTRA71.PLOT,
          /* Slope_aspect */
            (PUT(SLOPEEXTRA71.Slope_aspect, 3.0) ) AS Slope_aspect,
          /* PD_Code */
            (INPUT(SLOPEEXTRA71.PD_Code, 1.0) ) AS PD_Code,
          /* Year */
            ("1971") AS Year
      FROM WORK.SLOPEEXTRA71 SLOPEEXTRA71;
QUIT;
```

 # APPEND ADDTIONAL VALUES FOR SLOPE AND ASPECT

```
/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.allslope71 AS
SELECT * FROM WORK.QUERY_FOR_SLOPEEXTRA71
 OUTER UNION CORR
SELECT * FROM WORK.SLOPECODE571
;
Quit;

/*   START OF NODE: Query for allslope71   */

PROC SQL;
   CREATE TABLE WORK.Query_for_allslope71 AS
   SELECT ALLSLOPE71.SITE,
          ALLSLOPE71.PLOT,
          /* Slope_Aspect */
            (Input(ALLSLOPE71.Slope_aspect, 3.0) ) AS Slope_Aspect,
          ALLSLOPE71.PD_Code,
          ALLSLOPE71.Year
      FROM WORK.ALLSLOPE71 ALLSLOPE71;
QUIT;
```

```
/*   START OF NODE: Append Table   */
%LET _CLIENTTASKLABEL='Append Table';
PROC SQL;
CREATE TABLE WORK.APPEND_TABLE_0000 AS
SELECT * FROM WORK.QUERY_FOR_ALLSLOPE71
 OUTER UNION CORR
SELECT * FROM WORK.SLOPECODE503
;
Quit;


/*   START OF NODE: Query for APPEND_TABLE_0000   */

PROC SQL;
   CREATE TABLE WORK.Query_for_APPEND_TABLE_0000 AS
   SELECT APPEND_TABLE_0000.SITE,
          APPEND_TABLE_0000.PLOT,
          APPEND_TABLE_0000.Slope_Aspect AS Slope,
          APPEND_TABLE_0000.PD_Code,
          APPEND_TABLE_0000.Year
      FROM WORK.APPEND_TABLE_0000 APPEND_TABLE_0000;
QUIT;

/*   START OF NODE: Query1 for QUERY7867   */

PROC SQL;
   CREATE TABLE WORK.aspectextra71 AS
   SELECT QUERY7867.SITE,
          QUERY7867.PLOT,
          QUERY7867.PD_code AS Slope_aspect,
          QUERY7867.Slope_aspect AS PD_Code
      FROM WORK.QUERY7867 QUERY7867
      WHERE QUERY7867.Slope_aspect = '6';
QUIT;

/*   START OF NODE: Query for aspectextra71   */

PROC SQL;
   CREATE TABLE WORK.Query_for_aspectextra71 AS
   SELECT ASPECTEXTRA71.SITE,
          ASPECTEXTRA71.PLOT,
          /* Slope_aspect */
            (Put(ASPECTEXTRA71.Slope_aspect ,3.0)) AS Slope_aspect,
          /* PD_code */
            (Input(ASPECTEXTRA71.PD_Code, 1.0) ) AS PD_code,
          /* Year */
            ("1971") AS Year
      FROM WORK.ASPECTEXTRA71 ASPECTEXTRA71;
QUIT;

/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.all_aspect71 AS
SELECT * FROM WORK.QUERY_FOR_ASPECTEXTRA71
 OUTER UNION CORR
SELECT * FROM WORK.ASPECT71
;
Quit;

/*   START OF NODE: Query for all_aspect71   */

PROC SQL;
   CREATE TABLE WORK.Query_for_all_aspect71 AS
   SELECT ALL_ASPECT71.SITE,
          ALL_ASPECT71.PLOT,
          /* Slope_aspect */
```

```
            (Input(ALL_ASPECT71.Slope_aspect,3.0) ) AS Slope_aspect,
          ALL_ASPECT71.PD_code,
          ALL_ASPECT71.Year
      FROM WORK.ALL_ASPECT71 ALL_ASPECT71;
QUIT;

/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.APPEND_TABLE_0001 AS
SELECT * FROM WORK.QUERY_FOR_ALL_ASPECT71
 OUTER UNION CORR
SELECT * FROM WORK.ASPECT03
;
Quit;

/*   START OF NODE: Query for APPEND_TABLE_0001   */

PROC SQL;
   CREATE TABLE WORK.Query_for_APPEND_TABLE_0001 AS
   SELECT APPEND_TABLE_0001.SITE,
          APPEND_TABLE_0001.PLOT,
          APPEND_TABLE_0001.Slope_aspect AS Aspect,
          APPEND_TABLE_0001.PD_code,
          APPEND_TABLE_0001.Year
      FROM WORK.APPEND_TABLE_0001 APPEND_TABLE_0001;
QUIT;

/*   START OF NODE: Query for Query_for_APPEND_TABLE_0000   */

PROC SQL;
   CREATE TABLE WORK.Plot_slope_aspect AS
   SELECT QUERY_FOR_APPEND_TABLE_0000.SITE,
          QUERY_FOR_APPEND_TABLE_0000.PLOT,
          QUERY_FOR_APPEND_TABLE_0000.Slope AS 'Slope 'n,
          QUERY_FOR_APPEND_TABLE_0001.Aspect AS 'Aspect 'n,
          QUERY_FOR_APPEND_TABLE_0000.Year
      FROM WORK.QUERY_FOR_APPEND_TABLE_0000 QUERY_FOR_APPEND_TABLE_0000,
WORK.QUERY_FOR_APPEND_TABLE_0001
          QUERY_FOR_APPEND_TABLE_0001
      WHERE (QUERY_FOR_APPEND_TABLE_0000.SITE = QUERY_FOR_APPEND_TABLE_0001.SITE
AND QUERY_FOR_APPEND_TABLE_0000.PLOT =
           QUERY_FOR_APPEND_TABLE_0001.PLOT AND QUERY_FOR_APPEND_TABLE_0000.Year =
QUERY_FOR_APPEND_TABLE_0001.Year)
      ORDER BY QUERY_FOR_APPEND_TABLE_0000.SITE,
               QUERY_FOR_APPEND_TABLE_0000.PLOT,
               QUERY_FOR_APPEND_TABLE_0000.Year;
QUIT;
```

 # JOIN DESCRIPTIONS TO PLOT CODES 2003

```
/*   START OF NODE: Query for WORK.PLOTS03   */

PROC SQL;
   CREATE TABLE WORK.QUERY3117 AS
   SELECT PLOTS03.SITE LABEL="SITE",
          PLOTS03.PLOT LABEL="PLOT",
          PLOTS03.Notes LABEL="Notes",
          PLOTS03.PD_code LABEL="PD_code",
          PLOTLIB.Description LABEL="Description",
          /* Year */
            (2003) AS Year
      FROM WORK.PLOTS03 PLOTS03
           INNER JOIN WORK.PLOTLIB PLOTLIB ON (PLOTS03.PD_code = PLOTLIB.'Plot
code'n);
QUIT;
```

 # JOIN DESCRIPTIONS TO PLOT CODES 1971

```
/*   START OF NODE: Query for WORK.PLOTS71   */

PROC SQL;
   CREATE TABLE WORK.QUERY817 AS
   SELECT PLOTS71.SITE LABEL="SITE",
          PLOTS71.PLOT LABEL="PLOT",
          PLOTS71.Notes LABEL="Notes",
          PLOTS71.PD_code LABEL="PD_code",
          PLOTLIB.Description LABEL="Description",
          /* Year */
            (1971) AS Year
      FROM WORK.PLOTS71 PLOTS71
           INNER JOIN WORK.PLOTLIB PLOTLIB ON (PLOTS71.PD_code = PLOTLIB.'Plot
code'n);
QUIT;

/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.Plot_codes AS
SELECT * FROM WORK.QUERY3117
 OUTER UNION CORR
SELECT * FROM WORK.QUERY817
;
Quit;
```

 # IMPORT PLOT DISTURBANCE INFORMATION INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.IMPW_0007;
    LENGTH
        SITE               8
        PLOT               8
        NOTES            $ 24
        Built            $ 7
        Transport        $ 11
        Agriculture      $ 13
        Quarrying        $ 11
        'No access'n     $ 11 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
4e2d1da621574ee78c5d8786532e65f5.txt'
        LRECL=37
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        NOTES            :
        Built            :
        Transport        :
        Agriculture      :
        Quarrying        :
        'No access'n     :  ;
RUN;
```

 # APPEND DATA AND CREATE FINAL PLOT INFORMATION TABLE

```
/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.APPEND_TABLE_0006 AS
SELECT * FROM WORK.PLOT_CODES
```

```
 OUTER UNION CORR
SELECT * FROM WORK.IMPW_0007
;
Quit;


/*   START OF NODE: Query for APPEND_TABLE_0006   */

PROC SQL;
   CREATE TABLE WORK.Query_for_APPEND_TABLE_0006 AS
   SELECT DISTINCT APPEND_TABLE_0006.SITE,
          APPEND_TABLE_0006.PLOT,
          APPEND_TABLE_0006.Notes,
          APPEND_TABLE_0006.PD_code,
          APPEND_TABLE_0006.Description,
          APPEND_TABLE_0006.Year,
          /* Field_sheet */
            ("Plot description") AS Field_sheet,
          /* Reason_for_nosurvey */
            (CASE  WHEN APPEND_TABLE_0006.Built ="y" THEN "Built" WHEN
APPEND_TABLE_0006.Transport = "y"  THEN
            "Transport" WHEN APPEND_TABLE_0006.Agriculture = "y" THEN "Agriculture"
WHEN APPEND_TABLE_0006.Quarrying  =
            "y" THEN "Quarrying" WHEN APPEND_TABLE_0006.'No access'n  = "y" THEN
"No access"  ELSE "" END ) AS
            Reason_for_nosurvey
      FROM WORK.APPEND_TABLE_0006 APPEND_TABLE_0006
      GROUP BY APPEND_TABLE_0006.SITE,
               APPEND_TABLE_0006.PLOT,
               APPEND_TABLE_0006.Notes,
               APPEND_TABLE_0006.PD_code,
               APPEND_TABLE_0006.Description,
               APPEND_TABLE_0006.Year;
QUIT;
```

GBW4: Code to transform Woodland Survey soil information data to new schema
(Final table: Woodlands_Survey_Soil_Data_1971_2001.csv)

 # IMPORT SOIL PH DATA INTO SAS PROJECT

```
DATA WORK.PH;
    LENGTH
        SITE               8
        PLOT               8
        pH1971             8
        'pH resurvey'n     8
        pHchange           8
        pH1971_QA_repeats   8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
3bd8a7bd0c194469ad442104df6a38f6.txt'
        LRECL=41
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        pH1971           : BEST32.
        'pH resurvey'n   : BEST32.
        pHchange         : BEST32.
        pH1971_QA_repeats : BEST32. ;
```

```
RUN;
```

 # IMPORT SOIL ORGANIC MATTER DATA INTO SAS PROJECT

```
DATA WORK.SOM;
    LENGTH
        SITE              8
        PLOT              8
        SOM1971           8
        'SOM resurvey'n   8
        'SOM change'n     8
        SOM1971_QArepeats   8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
44bc493ec0a94c308bba4f508973a5fb.txt'
        LRECL=46
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        SOM1971          : BEST32.
        'SOM resurvey'n  : BEST32.
        'SOM change'n    : BEST32.
        SOM1971_QArepeats : BEST32. ;
RUN;
```

 # IMPORT SOIL GROUP DATA INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.IMPW;
    LENGTH
        SITE              8
        PLOT              8
        MSG               8
        'major soil group'n $ 18
        SG                8
        'soil group'n    $ 12
        'soil subgroup'n $ 15
        SSG               8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
cc5222faef824c51ba9c102bb6386fd5.txt'
        LRECL=89
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        MSG              : BEST32.
        'major soil group'n :
        SG               : BEST32.
        'soil group'n    :
        'soil subgroup'n :
        SSG              : BEST32. ;
RUN;
```

 # JOIN SOILS DATA SETS INTO ONE TO CREATE FINAL TABLE

```
/*   START OF NODE: Query for WORK.PH   */

PROC SQL;
   CREATE TABLE WORK.QUERY3350 AS
```

```
   SELECT DISTINCT PH.SITE LABEL="SITE",
          PH.PLOT LABEL="PLOT",
          PH.pH1971 LABEL="pH1971",
          PH.'pH resurvey'n LABEL="pH resurvey" AS pH2003,
          PH.pH1971_QA_repeats LABEL="pH1971_QA_repeats",
          SOM.SOM1971 LABEL="SOM1971",
          SOM.'SOM resurvey'n LABEL="SOM resurvey" AS SOM2003,
          SOM.SOM1971_QArepeats LABEL="SOM1971_QArepeats",
          IMPW.MSG LABEL="MSG",
          IMPW.SG LABEL="SG",
          IMPW.SSG LABEL="SSG",
          IMPW.'major soil group'n LABEL="major soil group" AS Major_Soil_Group,
          IMPW.'soil group'n LABEL="soil group" AS Soil_Group,
          IMPW.'soil subgroup'n LABEL="soil subgroup" AS Soil_subgroup
      FROM WORK.PH PH
           FULL JOIN WORK.SOM SOM ON (PH.SITE = SOM.SITE) AND (PH.PLOT = SOM.PLOT)
           FULL JOIN WORK.IMPW IMPW ON (PH.SITE = IMPW.SITE) AND (PH.PLOT =
IMPW.PLOT)
      GROUP BY PH.SITE,
               PH.PLOT,
               PH.pH1971,
               PH.'pH resurvey'n,
               PH.pH1971_QA_repeats,
               SOM.SOM1971,
               SOM.'SOM resurvey'n,
               SOM.SOM1971_QArepeats,
               IMPW.MSG,
               IMPW.SG,
               IMPW.SSG,
               IMPW.'major soil group'n,
               IMPW.'soil group'n,
               IMPW.'soil subgroup'n
      ORDER BY PH.SITE,
               PH.PLOT;
QUIT;
```

GBW5: Code to incorporate Avery soil codes into Woodlands data

(Final table: Woodlands_Survey_Soil_Data_1971_2001.csv)

#IMPORT TABLE TO SAS: Woodland Survey soil data

```
/*   START OF NODE: Import Data (Woodland_Survey_Soil_Data.csv)   */

DATA WORK.Woodland_Survey_Soil_Data;
    LENGTH
        SITE               8
        PLOT               8
        pH1971             8
        pH2003             8
        pH1971_QA_repeats   8
        SOM1971            8
        SOM2003            8
        SOM1971_QArepeats   8
        MSG                8
        SG                 8
        SSG                8
        Major_Soil_Group $ 18
        Soil_Group       $ 12
        Soil_subgroup    $ 15 ;
    FORMAT
        SITE             BEST3.
        PLOT             BEST2.
        pH1971           BEST3.
```

```
        pH2003           BEST5.
        pH1971_QA_repeats BEST5.
        SOM1971          BEST6.
        SOM2003          BEST8.
        SOM1971_QArepeats BEST6.
        MSG              BEST2.
        SG               BEST4.
        SSG              BEST5.
        Major_Soil_Group $CHAR18.
        Soil_Group       $CHAR12.
        Soil_subgroup    $CHAR15. ;
    INFORMAT
        SITE             BEST3.
        PLOT             BEST2.
        pH1971           BEST3.
        pH2003           BEST5.
        pH1971_QA_repeats BEST5.
        SOM1971          BEST6.
        SOM2003          BEST8.
        SOM1971_QArepeats BEST6.
        MSG              BEST2.
        SG               BEST4.
        SSG              BEST5.
        Major_Soil_Group $CHAR18.
        Soil_Group       $CHAR12.
        Soil_subgroup    $CHAR15. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG12312\Woodland_Survey_Soil_Data-
7d5e3544933249b1983f9cfe4bca5251.txt'
        LRECL=130
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : ?? BEST3.
        PLOT             : ?? BEST2.
        pH1971           : ?? COMMA3.
        pH2003           : ?? COMMA5.
        pH1971_QA_repeats : ?? COMMA5.
        SOM1971          : ?? COMMA6.
        SOM2003          : ?? COMMA8.
        SOM1971_QArepeats : ?? COMMA6.
        MSG              : ?? BEST2.
        SG               : ?? COMMA4.
        SSG              : ?? COMMA5.
        Major_Soil_Group : $CHAR18.
        Soil_Group       : $CHAR12.
        Soil_subgroup    : $CHAR15. ;
RUN;
```

#JOIN TABLE: Avery codes to soil data

```
/*   START OF NODE: Query Builder   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_WOODLAND_SURVEY_SOIL_D(label="QUERY_FOR_WOODLAND_SURVEY_SOIL_DATA")
AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.pH1971,
          t1.pH2003,
          t1.pH1971_QA_repeats,
          t1.SOM1971,
          t1.SOM2003,
          t1.SOM1971_QArepeats,
          t1.MSG,
```

```
          t1.SG,
          t1.SSG,
          t1.Soil_subgroup LENGTH=8
      FROM WORK.WOODLAND_SURVEY_SOIL_DATA t1
           LEFT JOIN CS_SOIL.CS_SOILS_AVERY_CODES t2 ON (t1.SSG =
t2.SUB_GROUP_CODE);
QUIT;
```

#IMPORT TABLE: Avery codes

```
/*   START OF NODE: Import Data (CS_SOILS_AVERY_CODES.csv)   */


DATA WORK.CS_SOILS_AVERY_CODES;
    LENGTH
        MAJOR_GROUP_CODE   8
        MAJOR_GROUP_DESCRIPTION $ 24
        GROUP_CODE         8
        GROUP_DESCRIPTION $ 32
        SUB_GROUP_CODE     8
        SUB_GROUP_DESCRIPTION $ 42 ;
    FORMAT
        MAJOR_GROUP_CODE BEST2.
        MAJOR_GROUP_DESCRIPTION $CHAR24.
        GROUP_CODE       BEST4.
        GROUP_DESCRIPTION $CHAR32.
        SUB_GROUP_CODE   BEST5.
        SUB_GROUP_DESCRIPTION $CHAR42. ;
    INFORMAT
        MAJOR_GROUP_CODE BEST2.
        MAJOR_GROUP_DESCRIPTION $CHAR24.
        GROUP_CODE       BEST4.
        GROUP_DESCRIPTION $CHAR32.
        SUB_GROUP_CODE   BEST5.
        SUB_GROUP_DESCRIPTION $CHAR42. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG12312\CS_SOILS_AVERY_CODES-
9d9d75f6c60944368e76addd39c8a380.txt'
        LRECL=100
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        MAJOR_GROUP_CODE : ?? BEST2.
        MAJOR_GROUP_DESCRIPTION : $CHAR24.
        GROUP_CODE       : ?? COMMA4.
        GROUP_DESCRIPTION : $CHAR32.
        SUB_GROUP_CODE   : ?? COMMA5.
        SUB_GROUP_DESCRIPTION : $CHAR42. ;
RUN;
```

#OUTPUT TABLE: QUERY_FOR_WOODLAND_SURVEY_FINAL

```
/*   START OF NODE: Query Builder1   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_WOODLAND_SURVEY_FINAL(label="QUERY_FOR_WOODLAND_SURVEY_SOIL_DATA")
AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.pH1971,
          t1.pH2003,
          t1.pH1971_QA_repeats,
          t1.SOM1971,
          t1.SOM2003,
          t1.SOM1971_QArepeats,
```

```
          t1.MSG,
          t1.SG,
          t1.SSG,
          t2.SUB_GROUP_DESCRIPTION,
          t2.GROUP_DESCRIPTION,
          t2.MAJOR_GROUP_DESCRIPTION
      FROM WORK.WOODLAND_SURVEY_SOIL_DATA t1
           LEFT JOIN WORK.CS_SOILS_AVERY_CODES t2 ON (t1.SSG = t2.SUB_GROUP_CODE);
QUIT;
```

GBW6: Code to transform Woodland Survey site data to new schema

(Final table: Woodlands_Survey_Site_Information_1971_2001.csv)

 # IMPORT SITE LIST (2003) INTO SAS PROJECT

```
DATA WORK.SITES03;
    LENGTH
        SITE               8
        SD_code            8
        Notes            $ 7
        Recorders        $ 11 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
a43dc7e09e5f44b0a6684d0a63506706.txt'
        LRECL=56
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        SD_code          : BEST32.
        Notes            :
        Recorders        :  ;
RUN;

/*   START OF NODE: Query for WORK.SITES03   */

PROC SQL;
   CREATE TABLE WORK.QUERY1033 AS
   SELECT SITES03.SITE LABEL="SITE",
          SITES03.SD_code LABEL="SD_code",
          SITES03.Notes LABEL="Notes",
          SITES03.Recorders LABEL="Recorders",
          /* Year */
            ("2003") AS Year
      FROM WORK.SITES03 SITES03;
QUIT;
```

 # IMPORT SITE LIST (1971) INTO SAS PROJECT

```
DATA WORK.SITES71;
    LENGTH
        SITE               8
        SD_code            8
        Notes            $ 7 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
be8b3d51647c4566a79976a4114f857b.txt'
        LRECL=72
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
```

```
        SD_code          : BEST32.
        Notes            :  ;
RUN;

/*   START OF NODE: Query for WORK.SITES71   */

PROC SQL;
   CREATE TABLE WORK.QUERY6572 AS
   SELECT SITES71.SITE LABEL="SITE",
          SITES71.SD_code LABEL="SD_code",
          SITES71.Notes LABEL="Notes",
          /* Year */
            ("1971") AS Year
      FROM WORK.SITES71 SITES71;
QUIT;

/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.Append_Table AS
SELECT * FROM WORK.QUERY1033
 OUTER UNION CORR
SELECT * FROM WORK.QUERY6572
;
Quit;
```

 # IMPORT SITE CODES INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Bunce_survey_71-03.mdb)   */

DATA WORK.SITE_CODES;
    LENGTH
        SD_code            8
        Description      $ 16
        Code_group       $ 12
        Score              8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
a648980b1a764813979b42df9438701b.txt'
        LRECL=51
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SD_code          : BEST32.
        Description      :
        Code_group       :
        Score            : BEST32. ;
RUN;
```

 # IMPORT SITE SURVEYORS INTO SAS PROJECT

```
DATA WORK.IMPW_0008;
    LENGTH
        'Site  no'n        8
        '1971 Recorder'n $ 13
        'Site 'n           8
        '2002Recorder'n  $ 15 ;
    INFORMAT
        'Site  no'n      COMMA32.
        '1971 Recorder'n $13.
        'Site 'n         COMMA32.
        '2002Recorder'n  $15. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\all sites recorders final-
139823276894458aa83446d65434e83d.txt'
        LRECL=21
        ENCODING="WLATIN1"
```

```
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Site  no'n      : ?? COMMA32.
        '1971 Recorder'n : $13.
        'Site 'n         : ?? COMMA32.
        '2002Recorder'n  : $15. ;
RUN;
```

 # IMPORT PLOT AND SITE DESCRIPTIONS INTO SAS PROJECT

```
DATA WORK.DATES71;
    LENGTH
        ID                 8
        'Site  no'n        8
        Name             $ 10
        Date             $ 12 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Plot and site descriptions
database-a89557e6746b481ab430504aec159245.txt'
        LRECL=26
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        ID               : BEST32.
        'Site  no'n      : BEST32.
        Name             :
        Date             :  ;
RUN;
```

 # IMPORT PLOT AND SITE DESCRIPTIONS (MANAGEMENT DATA) INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Plot and site descriptions database.mdb)   */

DATA WORK.MGT;
    LENGTH
        ID                 8
        'SITE No'n         8
        USAGE            $ 20 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Plot and site descriptions
database-748f0a58b5274fc88f5999317e48a2b1.txt'
        LRECL=26
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        ID               : BEST32.
        'SITE No'n       : BEST32.
        USAGE            :  ;
RUN;
```

 # IMPORT PLOT SURVEY DATES INTO SAS PROJECT

```
/*   START OF NODE: Import Data (dates.xls)   */

DATA WORK.DATES03NEW;
    LENGTH
        SITE               8
        '2002/3'n          8 ;
    FORMAT
        '2002/3'n        DATETIME18.0 ;
    INFORMAT
```

```
        '2002/3'n        DATETIME18. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\dates-
abf7ff3c57ee4ffb842bc23cd8275508.txt'
        LRECL=14
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        '2002/3'n        : BEST32. ;
RUN;
```

 # JOIN ALL SITE TABLES INTO OUTPUT

```
/*   START OF NODE: Query for WORK.IMPW_0008   */

PROC SQL;
   CREATE TABLE WORK.QUERY4289 AS
   SELECT IMPW_0008.'Site  no'n LABEL="Site  no",
          IMPW_0008.'1971 Recorder'n LABEL="1971 Recorder" AS '''1971Recorder''
(1971 Recorder)'n,
          IMPW_0008.'2002Recorder'n LABEL="2002Recorder",
          DATES03NEW.'2002/3'n LABEL="2002/3" AS Date2003,
          DATES71.Date LABEL="Date" AS Date1971,
          MGT.USAGE LABEL="USAGE"
      FROM WORK.IMPW_0008 IMPW_0008, WORK.DATES71 DATES71, WORK.DATES03NEW
DATES03NEW, WORK.MGT MGT
      WHERE (IMPW_0008.'Site  no'n = DATES71.'Site  no'n AND IMPW_0008.'Site  no'n
= DATES03NEW.SITE AND IMPW_0008.
           'Site  no'n = MGT.'SITE No'n);
QUIT;
```

GBW7: Code to transform Woodland Survey site information data to new schema

(Final table: Woodlands_Survey_Site_Information_1971_2001.csv)

 # IMPORT SITE LOCATION DATA INTO SAS PROJECT

```
DATA WORK.LOCS;
    LENGTH
        SITE_NUM           8
        EASTING            8
        NORTHING           8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
a6e314ffefa04aaab0220ac7deff83ce.txt'
        LRECL=13
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE_NUM         : BEST32.
        EASTING          : BEST32.
        NORTHING         : BEST32. ;
RUN;
```

 # IMPORT SITE CLASSIFICATION DATA INTO SAS PROJECT

```
DATA WORK.IMPW_0000;
    LENGTH
```

```
        SITE              8
        'Site Name'n     $ 59
        'Survey team'n   $ 20
        Country          $ 10
        LC               $ 4
        Zone              8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
fff804f4760547de8b5d0472085b5833.txt'
        LRECL=101
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        'Site Name'n     :
        'Survey team'n   :
        Country          :
        LC               :
        Zone             : BEST32. ;
RUN;
```

 # JOIN SITE CLASSIFICATION DATA TO LOCATION INFORMATION

```
/*   START OF NODE: Query for WORK.IMPW_0000   */

PROC SQL;
   CREATE TABLE WORK.QUERY4512 AS
   SELECT DISTINCT IMPW_0000.SITE LABEL="SITE",
          IMPW_0000.'Site Name'n LABEL="Site Name" AS NAME,
          IMPW_0000.'Survey team'n LABEL="Survey team" AS 'SURVEY_TEAM_2003 (Survey
team)'n,
          IMPW_0000.Country LABEL="Country",
          LOCS.EASTING LABEL="EASTING",
          LOCS.NORTHING LABEL="NORTHING"
      FROM WORK.IMPW_0000 IMPW_0000
           INNER JOIN WORK.LOCS LOCS ON (IMPW_0000.SITE = LOCS.SITE_NUM)
      GROUP BY IMPW_0000.SITE,
               IMPW_0000.'Site Name'n,
               IMPW_0000.'Survey team'n,
               IMPW_0000.Country,
               LOCS.EASTING,
               LOCS.NORTHING;
QUIT;
```

 # IMPORT SITE SURVEY DATES DATA INTO SAS PROJECT

```
DATA WORK.DATES;
    LENGTH
        SITE              8
        '1971 DAY'n       8
        '1971 MONTH'n     8
        'RESURVEY DAY'n   8
        'RESURVEY MONTH'n   8
        'Day differences'n   8 ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13436\Bunce_survey_71-03-
1e471263ed3d403aa6d50f81bacc89af.txt'
        LRECL=16
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        '1971 DAY'n      : BEST32.
        '1971 MONTH'n    : BEST32.
```

```
        'RESURVEY DAY'n  : BEST32.
        'RESURVEY MONTH'n : BEST32.
        'Day differences'n : BEST32. ;
RUN;
```

 # JOIN DATES TO SITE CLASSIFICATION DATA AND LOCATION INFORMATION, NAME COLUMNS, CREATE FINAL TABLE

```
/*   START OF NODE: Query for QUERY4512   */

PROC SQL;
   CREATE TABLE WORK.Query_for_QUERY4512 AS
   SELECT DISTINCT QUERY4512.SITE,
          QUERY4512.NAME,
          QUERY4512.'SURVEY_TEAM_2003 (Survey team)'n,
          QUERY4512.Country,
          QUERY4512.EASTING,
          QUERY4512.NORTHING,
          DATES.SITE LABEL="SITE" AS SITE1,
          DATES.'1971 DAY'n LABEL="1971 DAY",
          DATES.'1971 MONTH'n LABEL="1971 MONTH",
          DATES.'RESURVEY DAY'n LABEL="RESURVEY DAY",
          DATES.'RESURVEY MONTH'n LABEL="RESURVEY MONTH",
          DATES.'Day differences'n LABEL="Day differences"
      FROM WORK.QUERY4512 QUERY4512
           INNER JOIN WORK.DATES DATES ON (QUERY4512.SITE = DATES.SITE)
      GROUP BY QUERY4512.SITE,
               QUERY4512.NAME,
               QUERY4512.'SURVEY_TEAM_2003 (Survey team)'n,
               QUERY4512.Country,
               QUERY4512.EASTING,
               QUERY4512.NORTHING,
               DATES.SITE,
               DATES.'1971 DAY'n,
               DATES.'1971 MONTH'n,
               DATES.'RESURVEY DAY'n,
               DATES.'RESURVEY MONTH'n,
               DATES.'Day differences'n;
QUIT;
```

GBW8: Code to transform Woodland Survey data to new schema

 # IMPORT PLOT DESCRIPTION DATA (2003) INTO SAS PROJECT

```
DATA WORK.Plot_descriptors00_03;
    LENGTH
        SITE               8
        PLOT               8
        Notes            $ 244
        PD_code            8 ;
    FORMAT
        SITE             BEST12.
        PLOT             BEST12.
        Notes            $CHAR244.
        PD_code          BEST12. ;
    INFORMAT
        SITE             BEST12.
        PLOT             BEST12.
        Notes            $CHAR244.
        PD_code          BEST12. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\Plot_descriptors00-03-
906528eae3d641178d458f28ff135863.txt'
        LRECL=251
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
```

```
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        Notes            : $CHAR244.
        PD_code          : BEST32. ;
RUN;
```

 # IMPORT PLOT DESCRIPTION DATA (1971) INTO SAS PROJECT

```
DATA WORK.Plot_descriptors1971;
    LENGTH
        SITE               8
        PLOT               8
        Notes            $ 44
        PD_code            8 ;
    FORMAT
        SITE             BEST12.
        PLOT             BEST12.
        Notes            $CHAR44.
        PD_code          BEST12. ;
    INFORMAT
        SITE             BEST12.
        PLOT             BEST12.
        Notes            $CHAR44.
        PD_code          BEST12. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\Plot_descriptors1971-
ff7088088ec241c492b3184b41b880df.txt'
        LRECL=52
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        Notes            : $CHAR44.
        PD_code          : BEST32. ;
RUN;
```

 # IMPORT PLOT DESCRIPTION CODE LOOKUP INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Plot description
code_LIBRARY.xlsx[Plot_description_code_LIBRARY])   */


DATA WORK.Plot_description_code_LIBRARY;
    LENGTH
        'Plot code'n       8
        Description      $ 20 ;
    FORMAT
        'Plot code'n     BEST12.
        Description      $CHAR20. ;
    INFORMAT
        'Plot code'n     BEST12.
        Description      $CHAR20. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\Plot description
code_LIBRARY-38d90da5ffb743ee9136a12097f1f409.txt'
        LRECL=22
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Plot code'n     : BEST32.
        Description      : $CHAR20. ;
RUN;
```

 # JOIN PLOT DESCRIPTION DATA, 1971 AND 2003 AND JOIN TO CODES

```
/*   START OF NODE: Query Builder1   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_PLOT_DESCRIPTORS00_03 AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.Notes,
          t1.PD_code,
          t2.Description,
          /* YEAR */
            (2001) AS YEAR
      FROM WORK.PLOT_DESCRIPTORS00_03 t1
           LEFT JOIN WORK.PLOT_DESCRIPTION_CODE_LIBRARY t2 ON (t1.PD_code =
t2.'Plot code'n);
QUIT;

/*   START OF NODE: Query Builder2   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_PLOT_DESCRIPTORS AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.Notes,
          t1.PD_code,
          t2.Description,
          /* YEAR */
            (1971) AS YEAR
      FROM WORK.PLOT_DESCRIPTORS1971 t1
           LEFT JOIN WORK.PLOT_DESCRIPTION_CODE_LIBRARY t2 ON (t1.PD_code =
t2.'Plot code'n);
QUIT;
```

 # IMPORT TABLE 'MISSING PLOTS' INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Missing_plots.xlsx[Missing_plots])   */

DATA WORK.Missing_plots;
    LENGTH
        SITE               8
        PLOT               8
        NOTES            $ 22
        Built            $ 1
        Transport        $ 1
        Agriculture      $ 1
        Quarrying        $ 1
        Noaccess         $ 1 ;
    FORMAT
        SITE             BEST12.
        PLOT             BEST12.
        NOTES            $CHAR22.
        Built            $CHAR1.
        Transport        $CHAR1.
        Agriculture      $CHAR1.
        Quarrying        $CHAR1.
        Noaccess         $CHAR1. ;
    INFORMAT
        SITE             BEST12.
        PLOT             BEST12.
        NOTES            $CHAR22.
        Built            $CHAR1.
        Transport        $CHAR1.
        Agriculture      $CHAR1.
        Quarrying        $CHAR1.
        Noaccess         $CHAR1. ;
```

```
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\Missing_plots-
53583d5c764c4532a1a94787a42acefb.txt'
        LRECL=37
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        PLOT             : BEST32.
        NOTES            : $CHAR22.
        Built            : $CHAR1.
        Transport        : $CHAR1.
        Agriculture      : $CHAR1.
        Quarrying        : $CHAR1.
        Noaccess         : $CHAR1. ;
RUN;
```

 # CATEGORISE REASONS FOR PLOT NOT BEING SURVEYED

```
/*   START OF NODE: Query Builder   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_MISSING_PLOTS AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.NOTES,
          /* Reason_nosurvey */
            (CASE
               WHEN t1.Built = "y"
               THEN "Built"
            WHEN t1.Transport= "y"
               THEN "Transport"
            WHEN t1.Agriculture= "y"
               THEN "Agriculture"
            WHEN t1.Quarrying= "y"
               THEN "Quarrying"
            WHEN t1.Noaccess= "y"
               THEN "No access"
               ELSE ""
            END) AS Reason_nosurvey,
          /* YEAR */
            (2001) AS YEAR
      FROM WORK.MISSING_PLOTS t1;
QUIT;
```

 # APPEND TABLES TOGETHER AND SORT

```
/*   START OF NODE: Append Table   */

PROC SQL;
CREATE TABLE WORK.Append_Table AS
SELECT * FROM WORK.QUERY_FOR_PLOT_DESCRIPTORS00_03
 OUTER UNION CORR
SELECT * FROM WORK.QUERY_FOR_MISSING_PLOTS
 OUTER UNION CORR
SELECT * FROM WORK.QUERY_FOR_PLOT_DESCRIPTORS
;
Quit;

PROC SQL;
   CREATE TABLE WORK.FILTER_FOR_APPEND_TABLE AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.Notes,
          t1.PD_code,
          t1.Description,
```

```
          t1.YEAR,
          t1.Reason_nosurvey
      FROM WORK.APPEND_TABLE t1
      ORDER BY t1.YEAR,
               t1.SITE,
               t1.PLOT,
               t1.PD_code;
QUIT;
```

 # ADD COLUMN TO DENOTE THE SPECIFIC RECORDING SHEET

```
/*   START OF NODE: Query Builder3   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_FILTER_FOR_APPEND_TABL(label="QUERY_FOR_FILTER_FOR_APPEND_TABLE") AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.Notes,
          t1.PD_code,
          t1.Description,
          t1.YEAR,
          t1.Reason_nosurvey,
          /* Field_sheet */
            ("Plot descriptions") AS Field_sheet
      FROM WORK.FILTER_FOR_APPEND_TABLE t1;
QUIT;
```

 # IMPORT SITE DESCRIPTION DATA (2003) INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Site_descriptors00-03.xlsx[Site_descriptors00_03])
*/

DATA WORK.Site_descriptors00_03;
    LENGTH
        SITE               8
        SD_code            8
        Notes            $ 100
        Recorders        $ 100 ;
    FORMAT
        SITE             BEST12.
        SD_code          BEST12.
        Notes            $CHAR100.
        Recorders        $CHAR100. ;
    INFORMAT
        SITE             BEST12.
        SD_code          BEST12.
        Notes            $CHAR100.
        Recorders        $CHAR100. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\Site_descriptors00-03-
e2ecb0ea167d4a3ba4c5d9cc496ee813.txt'
        LRECL=56
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE             : BEST32.
        SD_code          : BEST32.
        Notes            : $CHAR45.
        Recorders        : $CHAR100. ;
RUN;
```

 # IMPORT SITE DESCRIPTION DATA (1971) INTO SAS PROJECT

```
DATA WORK.Site_descriptors1971;
    LENGTH
```

```
        SITE              8
        SD_code           8
        Notes           $ 100 ;
    FORMAT
        SITE            BEST12.
        SD_code         BEST12.
        Notes           $CHAR100. ;
    INFORMAT
        SITE            BEST12.
        SD_code         BEST12.
        Notes           $CHAR100. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\Site_descriptors1971-
0eb8ac59ccc34136926d6a5622a9b92b.txt'
        LRECL=72
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SITE            : BEST32.
        SD_code         : BEST32.
        Notes           : $CHAR100. ;
RUN;
```

 # IMPORT SITE DESCRIPTION CODES INTO SAS PROJECT

```
/*   START OF NODE: Import Data (Site description
code_LIBRARY.xlsx[Site_description_code_LIBRARY])   */

DATA WORK.Site_description_code_LIBRARY;
    LENGTH
        SD_code           8
        Description     $ 43
        Code_group      $ 2
        Score             8 ;
    FORMAT
        SD_code         BEST12.
        Description     $CHAR43.
        Code_group      $CHAR2.
        Score           BEST12. ;
    INFORMAT
        SD_code         BEST12.
        Description     $CHAR43.
        Code_group      $CHAR2.
        Score           BEST12. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\Site description
code_LIBRARY-62d2860bf2ea4f2d8d425c46e1725175.txt'
        LRECL=51
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        SD_code         : BEST32.
        Description     : $CHAR43.
        Code_group      : $CHAR2.
        Score           : BEST32. ;
RUN;
```

 # JOIN SITE DESCRIPTION DATA, 1971 AND 2003 AND JOIN TO CODES

```
/*   START OF NODE: Query Builder6   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_SITE_DESCRIPTORS00_03 AS
   SELECT t1.SITE,
```

```
          t1.SD_code,
          t1.Notes,
          t1.Recorders,
          t2.Description,
          t2.Code_group,
          t2.Score,
          /* YEAR */
            (2001) AS YEAR
      FROM WORK.SITE_DESCRIPTORS00_03 t1
           LEFT JOIN WORK.SITE_DESCRIPTION_CODE_LIBRARY t2 ON (t1.SD_code =
t2.SD_code)
      WHERE t1.SITE NOT IN
           (
           53,
           23
           );
QUIT;

/*   START OF NODE: Query Builder8   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_SITE_DESCRIPTORS AS
   SELECT t1.SITE,
          t1.SD_code,
          t1.Notes,
          t2.Description,
          t2.Code_group,
          t2.Score,
          /* YEAR */
            (1971) AS YEAR
      FROM WORK.SITE_DESCRIPTORS1971 t1
           LEFT JOIN WORK.SITE_DESCRIPTION_CODE_LIBRARY t2 ON (t1.SD_code =
t2.SD_code);
QUIT;
```

 # IMPORT GROUP CODES INTO SAS PROJECT

```
/*   START OF NODE: Import Data (code_group_codes.xlsx[Sheet1])   */

DATA WORK.code_group_codes;
    LENGTH
        Code_group       $ 19
        Code             $ 2 ;
    FORMAT
        Code_group       $CHAR19.
        Code             $CHAR2. ;
    INFORMAT
        Code_group       $CHAR19.
        Code             $CHAR2. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\code_group_codes-
d883d08df668423ebeb44cd692ee1c53.txt'
        LRECL=22
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Code_group       : $CHAR19.
        Code             : $CHAR2. ;
RUN;
```

 # JOIN CODE GROUPINGS TO OUTPUT TABLE

```
/*   START OF NODE: Query Builder11   */

PROC SQL;
```

```
   CREATE TABLE WORK.QUERY_FOR_APPEND_TABLE_0000(label="QUERY_FOR_APPEND_TABLE") AS
   SELECT t1.SITE,
          t1.SD_code AS Code,
          t1.Description,
          t1.Notes,
          t1.Code_group,
          t2.Code_group AS Code_group_dsc,
          t1.Score,
          t1.YEAR,
          /* FIELD_SHEET */
            ("Site description") AS FIELD_SHEET
      FROM WORK.QUERY_FOR_SITE_DESCRIPTORS t1
           LEFT JOIN WORK.CODE_GROUP_CODES t2 ON (t1.Code_group = t2.Code)
      WHERE t1.SD_code NOT = 3
      ORDER BY t1.SITE,
               t1.SD_code;
QUIT;
```

 # IMPORT GROUP CODES INTO SAS PROJECT

```
/*   START OF NODE: Import Data (speciesListOriginal.xls[SPECIES4])   */

DATA WORK.speciesListOriginal;
    LENGTH
        Species          $ 33
        BRCcode            8
        COMMON_NAME      $ 20
        LATIN_NAME       $ 26
        check            $ 1 ;
    FORMAT
        Species          $CHAR33.
        BRCcode          BEST12.
        COMMON_NAME      $CHAR20.
        LATIN_NAME       $CHAR26.
        check            $CHAR1. ;
    INFORMAT
        Species          $CHAR33.
        BRCcode          BEST12.
        COMMON_NAME      $CHAR20.
        LATIN_NAME       $CHAR26.
        check            $CHAR1. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\speciesListOriginal-
3eaaf69f9d2a4e7fb3249fffbddf71e1.txt'
        LRECL=68
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Species          : $CHAR33.
        BRCcode          : BEST32.
        COMMON_NAME      : $CHAR20.
        LATIN_NAME       : $CHAR26.
        check            : $CHAR1. ;
RUN;
```

 # IMPORT VEGETATION HEIGHT DATA SETS INTO SAS PROJECT

```
DATA WORK._1_50_height;
    LENGTH
        'Site No'n         8
        'Plot No'n         8
        Quadrat            8
        type             $ 2
        Species          $ 14
        height             8 ;
    FORMAT
```

```
        'Site No'n       BEST2.
        'Plot No'n       BEST2.
        Quadrat          BEST1.
        type             $CHAR2.
        Species          $CHAR14.
        height           BEST4. ;
    INFORMAT
        'Site No'n       BEST2.
        'Plot No'n       BEST2.
        Quadrat          BEST1.
        type             $CHAR2.
        Species          $CHAR14.
        height           BEST4. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\1-50 height-
e1d9a12b92d246648b81b280f419433b.txt'
        LRECL=30
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Site No'n       : ?? BEST2.
        'Plot No'n       : ?? BEST2.
        Quadrat          : ?? BEST1.
        type             : $CHAR2.
        Species          : $CHAR14.
        height           : ?? COMMA4. ;
RUN;

DATA WORK._51_69_heights;
    LENGTH
        'Site No'n         8
        'Plot No'n         8
        Quadrat            8
        type             $ 2
        Species          $ 14
        height             8 ;
    FORMAT
        'Site No'n       BEST2.
        'Plot No'n       BEST2.
        Quadrat          BEST1.
        type             $CHAR2.
        Species          $CHAR14.
        height           BEST4. ;
    INFORMAT
        'Site No'n       BEST2.
        'Plot No'n       BEST2.
        Quadrat          BEST1.
        type             $CHAR2.
        Species          $CHAR14.
        height           BEST4. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\51-69 heights-
43fd265ebee84fd3940b24cfe71151db.txt'
        LRECL=29
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Site No'n       : ?? BEST2.
        'Plot No'n       : ?? BEST2.
        Quadrat          : ?? BEST1.
        type             : $CHAR2.
        Species          : $CHAR14.
        height           : ?? COMMA4. ;
RUN;
```

```
DATA WORK._70_103_heights;
    LENGTH
        'Site No'n          8
        'Plot No'n          8
        Quadrat             8
        type              $ 2
        Species           $ 14
        height              8 ;
    FORMAT
        'Site No'n        BEST3.
        'Plot No'n        BEST2.
        Quadrat           BEST1.
        type              $CHAR2.
        Species           $CHAR14.
        height            BEST4. ;
    INFORMAT
        'Site No'n        BEST3.
        'Plot No'n        BEST2.
        Quadrat           BEST1.
        type              $CHAR2.
        Species           $CHAR14.
        height            BEST4. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\70-103 heights-
eb60706050f94fc4aaef84993138f95a.txt'
        LRECL=29
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Site No'n        : ?? BEST3.
        'Plot No'n        : ?? BEST2.
        Quadrat           : ?? BEST1.
        type              : $CHAR2.
        Species           : $CHAR14.
        height            : ?? COMMA4. ;
RUN;
```

 # APPEND VEGETATION HEIGHT DATA

```
/*   START OF NODE: Append Table1   */

PROC SQL;
CREATE TABLE WORK.APPEND_TABLE_0000 AS
SELECT * FROM WORK._1_50_HEIGHT
 OUTER UNION CORR
SELECT * FROM WORK._51_69_HEIGHTS
 OUTER UNION CORR
SELECT * FROM WORK._70_103_HEIGHTS
;
Quit;
```

 # ADD HEIGHTS TO SPECIES AND JOIN TO CODE DESCRIPTIONS

```
/*   START OF NODE: Query Builder4   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_APPEND_TABLE AS
   SELECT t1.'Site No'n AS SITE,
          t1.'Plot No'n AS PLOT,
          t1.Quadrat,
          t1.type,
          t1.Species,
          t1.height,
          t2.BRCcode,
```

```
            t2.LATIN_NAME,
            t2.COMMON_NAME,
            /* YEAR */
              (1971) AS YEAR
        FROM WORK.APPEND_TABLE_0000 t1
             LEFT JOIN WORK.SPECIESLISTORIGINAL t2 ON (t1.Species = t2.Species)
        WHERE t1.height NOT IS MISSING;
QUIT;
```

 # IMPORT DBH DATA INTO SAS PROJECT

```
/*   START OF NODE: Import Data (FINAL DBH Assembly.xls[Copy])   */

DATA WORK.FINAL_DBH_Assembly;
    LENGTH
        'site no'n           8
        recorder           $ 7
        date               $ 8
        'plot no'n           8
        quarter              8
        'TR/SA/SH'n        $ 3
        species            $ 28
        'BRC code'n          8
        height             $ 8
        DBH1               $ 33
        DBH2               $ 3
        DBH3               $ 3
        DBH4               $ 3
        DBH5               $ 3
        DBH6               $ 3
        DBH7               $ 3
        DBH8               $ 3
        DBH9               $ 3
        DBH10              $ 3
        DBH11              $ 3
        DBH12              $ 3
        DBH13              $ 3
        DBH14              $ 3
        DBH15              $ 3
        DBH16              $ 3
        DBH17              $ 3
        DBH18              $ 3
        DBH19              $ 3
        DBH20              $ 3
        DBH21              $ 3
        DBH22              $ 3
        DBH23              $ 3
        DBH24              $ 3
        DBH25              $ 3
        DBH26              $ 3
        DBH27              $ 3
        DBH28              $ 3
        DBH29              $ 3
        DBH30              $ 3
        DBH31              $ 3
        DBH32              $ 3
        DBH33              $ 3
        DBH34              $ 3
        DBH35              $ 3
        DBH36              $ 3
        DBH37              $ 3
        DBH38              $ 3
        DBH39              $ 2
        DBH40              $ 2
        DBH41              $ 2
        DBH42              $ 2
        DBH43              $ 2
        DBH44              $ 2
```

```
        DBH45            $ 2
        DBH46            $ 2
        DBH47            $ 2
        DBH48            $ 2
        DBH49            $ 2
        DBH50            $ 3
        DBH51            $ 3
        DBH52            $ 3
        DBH53            $ 3
        DBH54            $ 3
        DBH55            $ 2
        DBH56            $ 2
        DBH57            $ 2
        DBH58            $ 2
        DBH59            $ 2
        DBH60            $ 2
        F70              $ 2
        F71              $ 2
        F72              $ 2
        F73              $ 2
        F74              $ 2
        F75              $ 2
        F76              $ 2
        F77              $ 2
        F78              $ 2
        F79              $ 2
        F80              $ 2
        F81              $ 2
        F82              $ 2
        F83              $ 2
        F84              $ 2
        F85              $ 2 ;
    FORMAT
        'site no'n       BEST12.
        recorder         $CHAR7.
        date             $CHAR8.
        'plot no'n       BEST12.
        quarter          BEST12.
        'TR/SA/SH'n      $CHAR3.
        species          $CHAR28.
        'BRC code'n      BEST12.
        height           $CHAR8.
        DBH1             $CHAR33.
        DBH2             $CHAR3.
        DBH3             $CHAR3.
        DBH4             $CHAR3.
        DBH5             $CHAR3.
        DBH6             $CHAR3.
        DBH7             $CHAR3.
        DBH8             $CHAR3.
        DBH9             $CHAR3.
        DBH10            $CHAR3.
        DBH11            $CHAR3.
        DBH12            $CHAR3.
        DBH13            $CHAR3.
        DBH14            $CHAR3.
        DBH15            $CHAR3.
        DBH16            $CHAR3.
        DBH17            $CHAR3.
        DBH18            $CHAR3.
        DBH19            $CHAR3.
        DBH20            $CHAR3.
        DBH21            $CHAR3.
        DBH22            $CHAR3.
        DBH23            $CHAR3.
        DBH24            $CHAR3.
        DBH25            $CHAR3.
        DBH26            $CHAR3.
```

```
    DBH27             $CHAR3.
    DBH28             $CHAR3.
    DBH29             $CHAR3.
    DBH30             $CHAR3.
    DBH31             $CHAR3.
    DBH32             $CHAR3.
    DBH33             $CHAR3.
    DBH34             $CHAR3.
    DBH35             $CHAR3.
    DBH36             $CHAR3.
    DBH37             $CHAR3.
    DBH38             $CHAR3.
    DBH39             $CHAR2.
    DBH40             $CHAR2.
    DBH41             $CHAR2.
    DBH42             $CHAR2.
    DBH43             $CHAR2.
    DBH44             $CHAR2.
    DBH45             $CHAR2.
    DBH46             $CHAR2.
    DBH47             $CHAR2.
    DBH48             $CHAR2.
    DBH49             $CHAR2.
    DBH50             $CHAR3.
    DBH51             $CHAR3.
    DBH52             $CHAR3.
    DBH53             $CHAR3.
    DBH54             $CHAR3.
    DBH55             $CHAR2.
    DBH56             $CHAR2.
    DBH57             $CHAR2.
    DBH58             $CHAR2.
    DBH59             $CHAR2.
    DBH60             $CHAR2.
    F70               $CHAR2.
    F71               $CHAR2.
    F72               $CHAR2.
    F73               $CHAR2.
    F74               $CHAR2.
    F75               $CHAR2.
    F76               $CHAR2.
    F77               $CHAR2.
    F78               $CHAR2.
    F79               $CHAR2.
    F80               $CHAR2.
    F81               $CHAR2.
    F82               $CHAR2.
    F83               $CHAR2.
    F84               $CHAR2.
    F85               $CHAR2. ;
INFORMAT
    'site no'n        BEST12.
    recorder          $CHAR7.
    date              $CHAR8.
    'plot no'n        BEST12.
    quarter           BEST12.
    'TR/SA/SH'n       $CHAR3.
    species           $CHAR28.
    'BRC code'n       BEST12.
    height            $CHAR8.
    DBH1              $CHAR33.
    DBH2              $CHAR3.
    DBH3              $CHAR3.
    DBH4              $CHAR3.
    DBH5              $CHAR3.
    DBH6              $CHAR3.
    DBH7              $CHAR3.
    DBH8              $CHAR3.
```

```
DBH9                $CHAR3.
DBH10               $CHAR3.
DBH11               $CHAR3.
DBH12               $CHAR3.
DBH13               $CHAR3.
DBH14               $CHAR3.
DBH15               $CHAR3.
DBH16               $CHAR3.
DBH17               $CHAR3.
DBH18               $CHAR3.
DBH19               $CHAR3.
DBH20               $CHAR3.
DBH21               $CHAR3.
DBH22               $CHAR3.
DBH23               $CHAR3.
DBH24               $CHAR3.
DBH25               $CHAR3.
DBH26               $CHAR3.
DBH27               $CHAR3.
DBH28               $CHAR3.
DBH29               $CHAR3.
DBH30               $CHAR3.
DBH31               $CHAR3.
DBH32               $CHAR3.
DBH33               $CHAR3.
DBH34               $CHAR3.
DBH35               $CHAR3.
DBH36               $CHAR3.
DBH37               $CHAR3.
DBH38               $CHAR3.
DBH39               $CHAR2.
DBH40               $CHAR2.
DBH41               $CHAR2.
DBH42               $CHAR2.
DBH43               $CHAR2.
DBH44               $CHAR2.
DBH45               $CHAR2.
DBH46               $CHAR2.
DBH47               $CHAR2.
DBH48               $CHAR2.
DBH49               $CHAR2.
DBH50               $CHAR3.
DBH51               $CHAR3.
DBH52               $CHAR3.
DBH53               $CHAR3.
DBH54               $CHAR3.
DBH55               $CHAR2.
DBH56               $CHAR2.
DBH57               $CHAR2.
DBH58               $CHAR2.
DBH59               $CHAR2.
DBH60               $CHAR2.
F70                 $CHAR2.
F71                 $CHAR2.
F72                 $CHAR2.
F73                 $CHAR2.
F74                 $CHAR2.
F75                 $CHAR2.
F76                 $CHAR2.
F77                 $CHAR2.
F78                 $CHAR2.
F79                 $CHAR2.
F80                 $CHAR2.
F81                 $CHAR2.
F82                 $CHAR2.
F83                 $CHAR2.
F84                 $CHAR2.
F85                 $CHAR2. ;
```

```
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG11236\FINAL DBH Assembly -
111bc04fcd95462293b6279ae112f570.txt'
        LRECL=252
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'site no'n       : BEST32.
        recorder         : $CHAR7.
        date             : $CHAR8.
        'plot no'n       : BEST32.
        quarter          : BEST32.
        'TR/SA/SH'n      : $CHAR3.
        species          : $CHAR28.
        'BRC code'n      : BEST32.
        height           : $CHAR8.
        DBH1             : $CHAR33.
        DBH2             : $CHAR3.
        DBH3             : $CHAR3.
        DBH4             : $CHAR3.
        DBH5             : $CHAR3.
        DBH6             : $CHAR3.
        DBH7             : $CHAR3.
        DBH8             : $CHAR3.
        DBH9             : $CHAR3.
        DBH10            : $CHAR3.
        DBH11            : $CHAR3.
        DBH12            : $CHAR3.
        DBH13            : $CHAR3.
        DBH14            : $CHAR3.
        DBH15            : $CHAR3.
        DBH16            : $CHAR3.
        DBH17            : $CHAR3.
        DBH18            : $CHAR3.
        DBH19            : $CHAR3.
        DBH20            : $CHAR3.
        DBH21            : $CHAR3.
        DBH22            : $CHAR3.
        DBH23            : $CHAR3.
        DBH24            : $CHAR3.
        DBH25            : $CHAR3.
        DBH26            : $CHAR3.
        DBH27            : $CHAR3.
        DBH28            : $CHAR3.
        DBH29            : $CHAR3.
        DBH30            : $CHAR3.
        DBH31            : $CHAR3.
        DBH32            : $CHAR3.
        DBH33            : $CHAR3.
        DBH34            : $CHAR3.
        DBH35            : $CHAR3.
        DBH36            : $CHAR3.
        DBH37            : $CHAR3.
        DBH38            : $CHAR3.
        DBH39            : $CHAR2.
        DBH40            : $CHAR2.
        DBH41            : $CHAR2.
        DBH42            : $CHAR2.
        DBH43            : $CHAR2.
        DBH44            : $CHAR2.
        DBH45            : $CHAR2.
        DBH46            : $CHAR2.
        DBH47            : $CHAR2.
        DBH48            : $CHAR2.
        DBH49            : $CHAR2.
        DBH50            : $CHAR3.
```

```
        DBH51            : $CHAR3.
        DBH52            : $CHAR3.
        DBH53            : $CHAR3.
        DBH54            : $CHAR3.
        DBH55            : $CHAR2.
        DBH56            : $CHAR2.
        DBH57            : $CHAR2.
        DBH58            : $CHAR2.
        DBH59            : $CHAR2.
        DBH60            : $CHAR2.
        F70              : $CHAR2.
        F71              : $CHAR2.
        F72              : $CHAR2.
        F73              : $CHAR2.
        F74              : $CHAR2.
        F75              : $CHAR2.
        F76              : $CHAR2.
        F77              : $CHAR2.
        F78              : $CHAR2.
        F79              : $CHAR2.
        F80              : $CHAR2.
        F81              : $CHAR2.
        F82              : $CHAR2.
        F83              : $CHAR2.
        F84              : $CHAR2.
        F85              : $CHAR2. ;
RUN;
```

 # FILTER DATA TO REMOVE ROWS WITH NO HEIGHTS

```
/*   START OF NODE: Query Builder5   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_FINAL_DBH_ASSEMBLY AS
   SELECT t1.'site no'n,
          t1.'plot no'n,
          t1.quarter,
          t1.species,
          t1.'BRC code'n,
          t1.height,
          /* YEAR */
            (2001) AS YEAR
      FROM WORK.FINAL_DBH_ASSEMBLY t1
      WHERE t1.height NOT IS MISSING;
QUIT;
```

 # JOIN DATA TO SPECIES CODES AND DESCRIPTIONS

```
/*   START OF NODE: Query Builder7   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_FINAL_DBH_ASSEMBL_0000(label="QUERY_FOR_FINAL_DBH_ASSEMBLY") AS
   SELECT t1.'site no'n AS SITE,
          t1.'plot no'n AS PLOT,
          t1.quarter AS QUADRAT,
          t1.species,
          t1.'BRC code'n AS BRCcode,
          t1.height AS height1,
          t1.YEAR,
          t2.COMMON_NAME,
          t2.BRC_NAMES AS LATIN_NAME,
          /* HEIGHT */
            (INPUT(t1.height, BEST12.0)) AS HEIGHT
      FROM WORK.QUERY_FOR_FINAL_DBH_ASSEMBLY t1
           LEFT JOIN CSVEG_RO.LUS_SP_LIB_AND_TRAITS t2 ON (t1.'BRC code'n =
t2.BRC_NUMBER);
QUIT;
```

 # APPEND DBH DATA TO SPECIES LIST

```
/*   START OF NODE: Append Table4   */

PROC SQL;
CREATE TABLE WORK.APPEND_TABLE_0003 AS
SELECT * FROM WORK.QUERY_FOR_APPEND_TABLE
 OUTER UNION CORR
SELECT * FROM WORK.QUERY_FOR_FINAL_DBH_ASSEMBL_0000
;
Quit;
```

 # SORT DATA

```
/*   START OF NODE: Query Builder9   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_APPEND_TABLE_0001(label="QUERY_FOR_APPEND_TABLE") AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.Quadrat,
          t1.Species,
          t1.BRCcode,
          t1.LATIN_NAME,
          t1.height,
          t1.YEAR
      FROM WORK.APPEND_TABLE_0003 t1
      ORDER BY t1.SITE,
               t1.PLOT;
QUIT;
```

 # CREATE FINAL PLOT TABLE

```
/*   START OF NODE: Query Builder13   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_WOODLAND_SURVEY_PLOT_D(label="QUERY_FOR_WOODLAND_SURVEY_PLOT_DESCS")
AS
   SELECT t1.SITE,
          t1.PLOT,
          t1.NOTES,
          t1.PD_CODE AS CODE,
          t1.DESCRIPTION,
          t1.YEAR,
          t1.REASON_NOSURVEY,
          t1.FIELD_SHEET
      FROM BUNCE.WOODLAND_SURVEY_PLOT_DESCS t1;
QUIT;
```

 # CREATE FINAL SITE TABLE

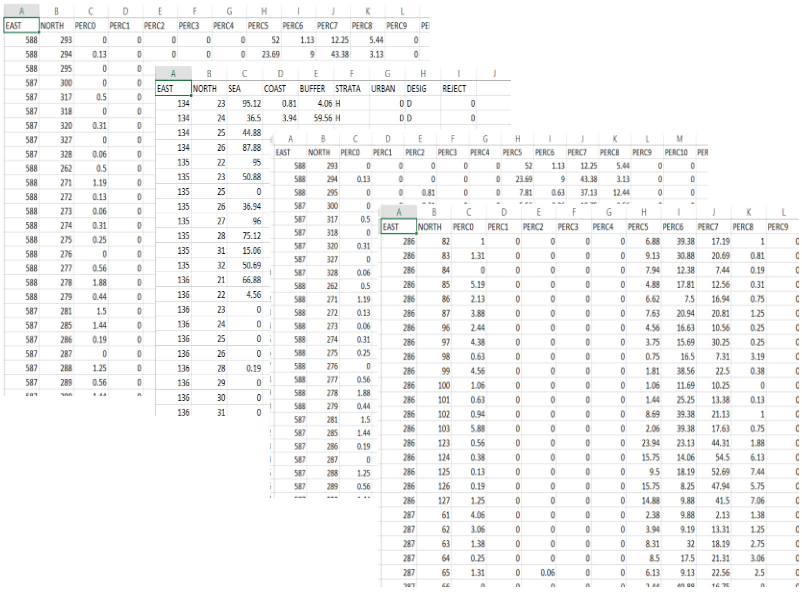
```
/*   START OF NODE: Query Builder12   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_WOODLAND_SURVEY_SITE_H(label="QUERY_FOR_WOODLAND_SURVEY_SITE_HEADERS
") AS
   SELECT t1.SITE_NUMBER AS SITE,
          t1.SITE_NAME,
          t1.EASTING,
          t1.NORTHING,
          t1.COUNTRY,
          t1.ITE_LAND_CLASS,
          t1.ITE_ENV_ZONE,
          t1.SURVEY_TEAM_2003,
          t1.RECORDERS1971,
```

```
        t1.DATE1971,
        t1.RECORDERS2003,
        t1.DATE2003,
        t1.USAGE,
        t1.SHAPE,
        /* FIELD_SHEET */
          ("Site header") AS FIELD_SHEET
    FROM BUNCE.WOODLAND_SURVEY_SITE_HEADERS t1;
QUIT;
```

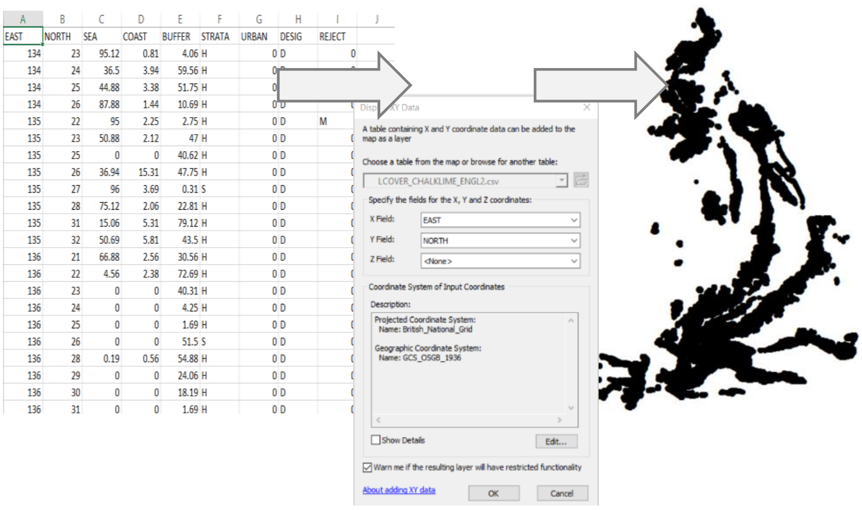
KHS1: Summary of procedure for creating mask/strata files for 'Key Habitat' Survey

1. Original data files located in old Oracle database and exported to csv files. Data structured as list of the bottom left corner of GB eastings and northings.

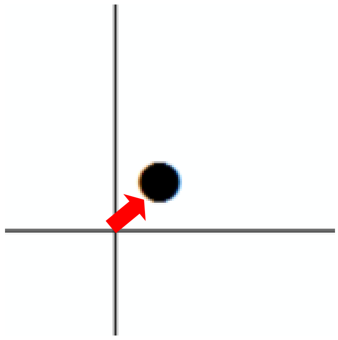
    - *LCOVER_COAST.csv*
    - *LCOVER_CHALKLIME_ENGL2.csv*
    - *LCOVER_LCLASS_HEATH_ENGL.csv*
    - *LCOVER_LCLASS_UPL_ENGL.csv*



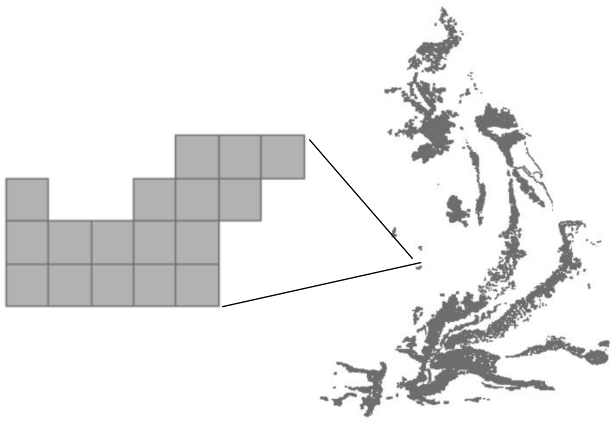
2. Files exported to spatial ESRI point shapefiles using the 'Display X,Y Data' tool in ArcGIS (and export results to shapefile).



3. Files filtered to return the correct number of rows, based on information available in Contract Reports
4. Files edited in ArcMap to be offset slightly, in order to be able to perform overlay with 1km grid of squares (otherwise points would be on exactly the grid intersections, not in the square areas).



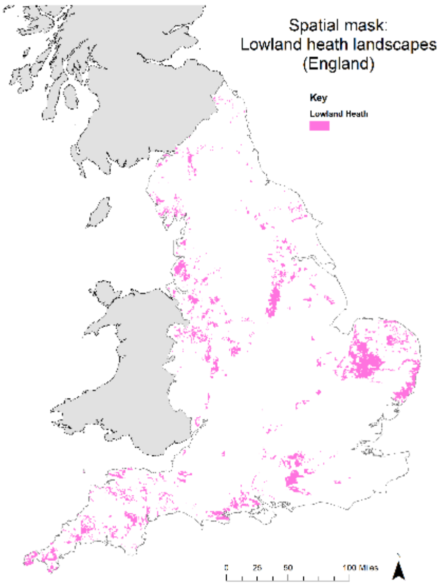
5. Points overlain with grids in ArcMap, using ‘Spatial Join’ tool creating 4 grids, one per mask type (calcareous, upland heath, lowland heath, coastal). This associated the data related to each point to an area (grid square).

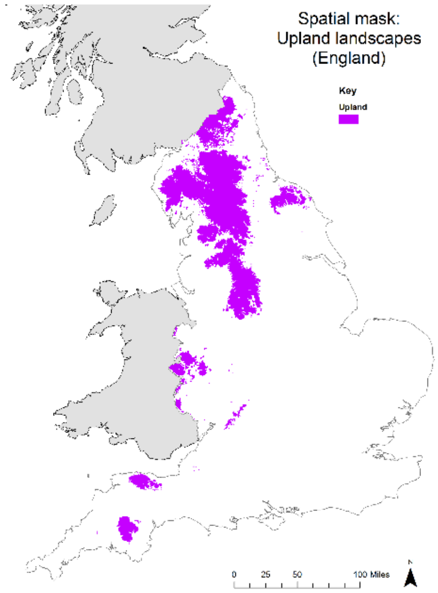


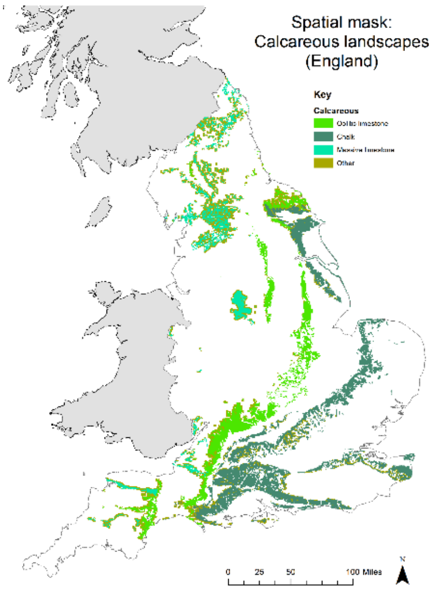
6. Data from each grid joined to form one dataset, with the 4 different data sets forming a separate column in the new data set:

BUNCE_SURVEYS.Key_Habitat_Masks

OBJECTID *	DOMGEOL	COAST	CALC	LO_HEATH	UP_HEATH	SHAPE *
1	0	1	0	0	0	Polygon
2	0	1	0	0	0	Polygon
3	0	1	0	0	0	Polygon
4	0	1	0	0	0	Polygon
5	0	1	0	0	0	Polygon
6	0	1	0	0	0	Polygon
7	0	1	0	0	0	Polygon
8	0	1	0	1	0	Polygon
9	0	1	0	1	0	Polygon
10	0	1	0	1	0	Polygon
11	0	1	0	0	0	Polygon
12	0	1	0	0	0	Polygon
13	0	1	0	1	0	Polygon
14	0	1	0	1	0	Polygon
15	0	1	0	1	0	Polygon
16	0	0	0	1	0	Polygon
17	0	1	0	1	0	Polygon
18	0	1	0	0	0	Polygon
19	0	1	0	0	0	Polygon
20	0	1	0	1	0	Polygon
21	0	0	0	1	0	Polygon
22	0	0	0	1	0	Polygon







7. Metadata document created based on known information regarding dataset (see Appendix iv for examples)

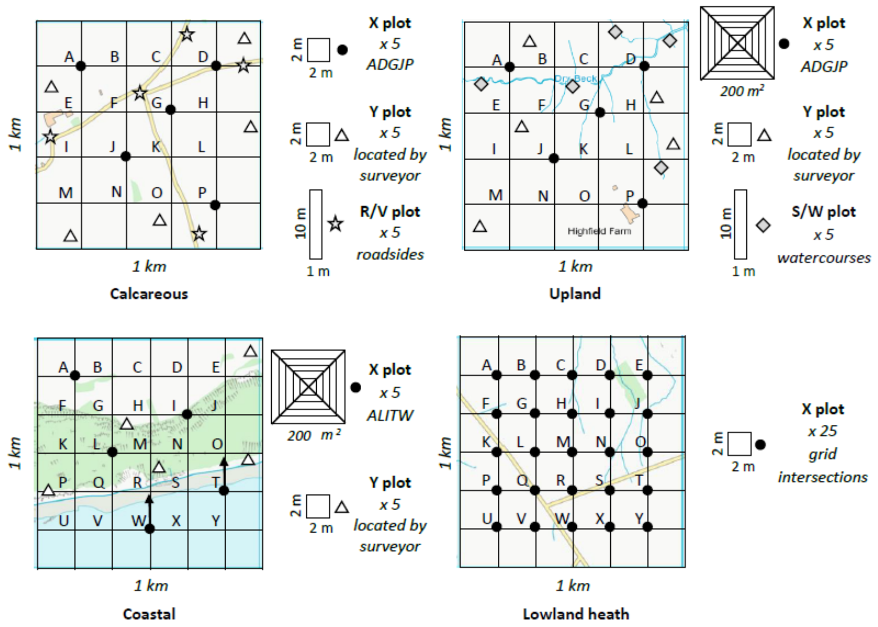
<u>Spatial masks for calcareous, coastal, upland and lowland heath habitats [Key Habitats 1992-93] (2017)</u>

Bunce, R.G.H, Parr, T.W., Ullyett, J., Hornung, M., Gerard, F., Bull, R., Cox, R., Brown, N.J.

Dataset Name:	***Spatial masks for calcareous, coastal, upland and lowland heath habitats [Key Habitats 1992-93]***
Dataset Background & Description:	Since 1978, a series of national surveys (Countryside Surveys) have been carried out by the Institute of Terrestrial Ecology/Centre for Ecology and Hydrology to gather data on the natural environment. The sampling framework for these surveys is not optimised to yield data on rarer or more specialised habitats. This Key Habitat survey was commissioned in the early 1990s to carry out additional survey work into habitats which were perceived to be under threat or which represented areas of concern to the Department for the Environment. The habitats were: Lowland heath, Chalk and limestone grasslands, Coasts, Uplands. For reporting purposes, the survey was supplemented by additional data, particularly for river valleys and waterside landscapes (lowlands), from Countryside Survey 1990. The masks provided by this dataset were used to define the population from which the 1km sampling sites were selected, and are thus existing or potential areas for the location of the habitats in question.
Geographic Coverage:	England
Time Period:	1992-93

KHS2: Procedure for recreating the location of vegetation plots in the Key Habitat Survey

The location of the recorded vegetation plots had to been digitised. However, it was known (based on information from the field sheets and handbooks, that plots were largely based on a grid system as follows:



Not to scale
Contains OS data © Crown copyright and database right (2018)

This meant that creating a file containing the location of the plots could largely be automated, negating the requirement for a lengthy digitising exercise. In order to do this, the following procedure was followed:

1. Grids of the appropriate size were generated for each landscape type, using the 'Create Fishnet' tool in ArcMap (e.g. 20x20m grids for the coastal and lowland heath sites, 25x25m grids for the calcareous and upland heath sites).
2. Based on the grid intersection sequences, the intersection points could be labelled with the identifiers A-P, or A-Y (again, depending on the landscape type).

POINTID	GRID_COD	POINT_X	POINT_Y	Code	Mapcode
1	8	489100.1	168900	1	A
2	8	489300.1	168900	2	B
3	8	489500.1	168900	3	C
4	8	489700.1	168900	4	D
5	8	489900.1	168900	5	E
6	8	489100.1	168700	6	F
7	8	489300.1	168700	7	G
8	8	489500.1	168700	8	H
9	8	489700.1	168700	9	I
10	8	489900.1	168700	10	J
11	8	489100.1	168500	11	K
12	8	489300.1	168500	12	L
13	8	489500.1	168500	13	M
14	8	489700.1	168500	14	N
15	8	489900.1	168500	15	O
16	8	489100.1	168300	16	P

The relevant points could then be matched to the digitised plot data, using the Site Square ID, and Plot Code (A-P or A-Y). Points not matching with recorded data were then removed from the dataset. A check was then made for plots still not having a location (by performing an outer join with the Recorded plot data and the points). Mismatches were then manually checked and digitised as necessary.

KHS3: Examples of SAS code to update species data to address quality issues in plot data

 # UPDATE PLOT SIZES IN SPECIES FILE WHERE NULL

```
proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES a
    set PLOT_SIZE = (select b.PLOT_SIZE from BUNCE.KEY_HABITAT_PLOT_HEADERS
b
            where a.series_num=b.series_num and a.rep_id=b.rep_id)
where a.PLOT_SIZE is null ;
quit;
```

 # UPDATE PLOT SIZES IN SPECIES FILE WHERE SITE IS AN9

```
/*proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES a
    set PLOT_SIZE = (select b.PLOT_SIZE from BUNCE.KEY_HABITAT_PLOT_HEADERS
b
            where a.series_num=b.series_num and a.rep_id=b.rep_id)
where a.SERIES_NUM = 'AN9' ;
quit;*/
```

 # UPDATE PLOT COVER VALUES IN SPECIES FILE WHERE SPECIES ARE PRESENT BUT COVER IS RECORDED AS NULL OR -1

```
proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES
    set PRES = 1
      where COVER < 0
      and PRES = 0;
     quit;

proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES
    set COVER = (COVER*-1)
      where COVER < 0 ;
     quit;

proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES
    set COVER2 = (COVER2*-1)
      where COVER2 < 0 ;
     quit;

proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES
    set COVER = COVER2
      where COVER NOT = COVER2 and PLOT_SIZE NOT='14.1x14.1' and COVER2 NOT
= 0 and COVER
and COVER2 NOT < COVER;
     quit;
```

 # UPDATE PLOT COVER VALUES IN SPECIES FILE WHERE SPECIES ARE BRYOPHYTES OR LICHENS/NOT BRYOPHYTES OR LICHENS

```
proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES
    set PRES = 1
where BRC_NAMES NOT in ('Total bryophyte', 'Total lichen')
and PRES = 0;
     quit;


proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES
    set PRES = 0
where BRC_NAMES in ('Total bryophyte', 'Total lichen')
;
     quit;
```

 # UPDATE PLOT COVER VALUES (NEST 2) TO 0 IN SPECIES FILE WHERE PLOT IS TOO SMALL TO HAVE A NEST 2

```
proc sql;
    update BUNCE.KEY_HABITAT_PLOT_SPECIES
    set COVER2 = 0
      where PLOT_SIZE NOT='14.1x14.1'
      ;
     quit;
```

KHS4: Key Habitats, code to transpose Land Cover Codes

 # IMPORT DATA INTO SAS PROJECT: LAND COVER RECORDED DATA

```
/*   START OF NODE: Import Data (LC_DATA.xlsx[QUERY_FOR_LC_CODES_WITH_PCO_000])

DATA WORK.LC_DATA;
    LENGTH
        Idno                8
        SP_GROUP            8
        HABT_GROUP          8
        SERIES_NUM        $ 5
        CODE_TYPE         $ 2
        GRIDCODE          $ 1
        FAB_POSITION        8
        HAB_TYPE          $ 13
        CODE                8
        DESCRIPTION       $ 59
        ADD_CODE          $ 3
        ADD_DESCRIPTION   $ 59
        POINT_X             8
        POINT_Y             8
        CAT               $ 7
        LUSE              $ 3 ;
    FORMAT
        Idno              BEST12.
        SP_GROUP          BEST12.
        HABT_GROUP        BEST12.
        SERIES_NUM        $CHAR5.
        CODE_TYPE         $CHAR2.
        GRIDCODE          $CHAR1.
        FAB_POSITION      BEST12.
        HAB_TYPE          $CHAR13.
```

```
        CODE             BEST12.
        DESCRIPTION      $CHAR59.
        ADD_CODE         $CHAR3.
        ADD_DESCRIPTION  $CHAR59.
        POINT_X          BEST12.
        POINT_Y          BEST12.
        CAT              $CHAR7.
        LUSE             $CHAR3. ;
    INFORMAT
        Idno             BEST12.
        SP_GROUP         BEST12.
        HABT_GROUP       BEST12.
        SERIES_NUM       $CHAR5.
        CODE_TYPE        $CHAR2.
        GRIDCODE         $CHAR1.
        FAB_POSITION     BEST12.
        HAB_TYPE         $CHAR13.
        CODE             BEST12.
        DESCRIPTION      $CHAR59.
        ADD_CODE         $CHAR3.
        ADD_DESCRIPTION  $CHAR59.
        POINT_X          BEST12.
        POINT_Y          BEST12.
        CAT              $CHAR7.
        LUSE             $CHAR3. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\LC_DATA-
72fc2581c8e24ae5bd9897a5761e313a.txt'
        LRECL=187
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Idno             : BEST32.
        SP_GROUP         : BEST32.
        HABT_GROUP       : BEST32.
        SERIES_NUM       : $CHAR5.
        CODE_TYPE        : $CHAR2.
        GRIDCODE         : $CHAR1.
        FAB_POSITION     : BEST32.
        HAB_TYPE         : $CHAR13.
        CODE             : BEST32.
        DESCRIPTION      : $CHAR59.
        ADD_CODE         : $CHAR3.
        ADD_DESCRIPTION  : $CHAR59.
        POINT_X          : BEST32.
        POINT_Y          : BEST32.
        CAT              : $CHAR7.
        LUSE             : $CHAR3. ;
RUN;
```

 # IMPORT DATA INTO SAS PROJECT: LAND COVER DESCRIPTION CODE LIST

```
/*   START OF NODE: Import Data (LC_code_list.xlsx[LC_code_list])   */

DATA WORK.LC_code_list;
    LENGTH
        CODE               8
        DESCRIPTION      $ 59
        PRIMARY          $ 2
        CAT              $ 7
        LUSE             $ 3 ;
    FORMAT
        CODE             BEST12.
        DESCRIPTION      $CHAR59.
        PRIMARY          $CHAR2.
        CAT              $CHAR7.
```

```
        LUSE             $CHAR3. ;
    INFORMAT
        CODE             BEST12.
        DESCRIPTION      $CHAR59.
        PRIMARY          $CHAR2.
        CAT              $CHAR7.
        LUSE             $CHAR3. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\LC_code_list-
dd15804538134ccaacaf912237319a22.txt'
        LRECL=72
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        CODE             : BEST32.
        DESCRIPTION      : $CHAR59.
        PRIMARY          : $CHAR2.
        CAT              : $CHAR7.
        LUSE             : $CHAR3. ;
RUN;
```

 # SELECT ALL RECORDED CODES AND DESCRIPTIONS FROM LAND COVER DATA

```
/*   START OF NODE: Query Builder33   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_LC_CODES_WITH_PCODE AS
   SELECT DISTINCT t1.CODE,
          t1.DESCRIPTION,
          t1.PRIMARY
      FROM WORK.LC_CODES_WITH_PCODE t1;
QUIT;
```

 # MATCH RECORDED CODES WITH DEFINITIVE CODE LIST, IN ORDER TO STANDARDISE RECORDED DESCRIPTIONS AND IDENTIFY NON-STANDARD ENTRIES

```
/*   START OF NODE: Query Builder34   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_LC_CODES_WITH_PCO_0000(label="QUERY_FOR_LC_CODES_WITH_PCODE") AS
   SELECT DISTINCT t1.Idno,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t2.CAT,
          t2.LUSE
      FROM WORK.LC_CODES_WITH_PCODE t1
           LEFT JOIN WORK.LC_CODE_LIST t2 ON (t1.CODE = t2.CODE) AND
(t1.DESCRIPTION = t2.DESCRIPTION)
      ORDER BY t1.SERIES_NUM,
               t1.CODE_TYPE,
               t1.GRIDCODE,
               t1.FAB_POSITION;
QUIT;
```

 # SPLIT DATA INTO THEMES: HABT (HABITATS)

```
/*   START OF NODE: Query Builder35   */

PROC SQL;
   CREATE TABLE WORK.LC_HABTS AS
   SELECT DISTINCT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'HABT';
QUIT;
```

 # SPLIT DATA INTO THEMES: PEAT

```
/*   START OF NODE: Query Builder36   */

PROC SQL;
   CREATE TABLE WORK.peat AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'PEAT';
QUIT;
```

 # SPLIT DATA INTO THEMES: AGUSE (AGRICULTURAL USE)

```
/*   START OF NODE: Query Builder37   */

PROC SQL;
   CREATE TABLE WORK.aguse AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
```

```
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'AGUSE';
QUIT;
```

 # SPLIT DATA INTO THEMES: NO (NUMBER, OF ITEMS IN PRIMARY CODE, HORSES)

```
/*   START OF NODE: Query Builder39   */

PROC SQL;
   CREATE TABLE WORK.no AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'NO';
QUIT;
```

 # REMOVE NEGATIVE VALUE SYMBOL FROM 'NO' (NUMBER) VALUES USING ABS FUNCTION

```
/*   START OF NODE: Query Builder57   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_NO AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          /* Calculation */
            (ABS(t1.CODE)) AS Calculation,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.NO t1;
QUIT;
```

 # JOIN NO (NUMBER) COLUMN FROM ABOVE TO ENTRIES OF 'HORSES'

```
/*   START OF NODE: Query Builder58   */

PROC SQL;
```

```
   CREATE TABLE WORK.aguse2 AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t2.Calculation AS Number,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.AGUSE t1
           LEFT JOIN WORK.QUERY_FOR_NO t2 ON (t1.HABT_GROUP = t2.HABT_GROUP AND
(t1.DESCRIPTION = "Horses"));
QUIT;
```

 # SPLIT DATA INTO THEMES: AGUSE (AGRICULTURAL USE)

```
/*   START OF NODE: Query Builder41   */

PROC SQL;
   CREATE TABLE WORK.forbs AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'FORBS';
QUIT;
```

 # SPLIT DATA INTO THEMES: USE

```
/*   START OF NODE: Query Builder42   */

PROC SQL;
   CREATE TABLE WORK.use AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
```

```
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'USE';
QUIT;
```

 # SPLIT DATA INTO THEMES: DESC (DESCRIPTION)

```
/*   START OF NODE: Query Builder44   */

PROC SQL;
   CREATE TABLE WORK.desc AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'DESC';
QUIT;
```

 # SPLIT DATA INTO THEMES: CANOPY

```
/*   START OF NODE: Query Builder45   */

PROC SQL;
   CREATE TABLE WORK.can AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'CANOPY';
QUIT;
```

 # SPLIT DATA INTO THEMES: MGT (MANAGEMENT)

```
/*   START OF NODE: Query Builder48   */

PROC SQL;
   CREATE TABLE WORK.mgt AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
```

```
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'MGT';
QUIT;
```

 # SPLIT DATA INTO THEMES: FOF (FORESTRY FEATURE)

```
/*   START OF NODE: Query Builder49   */

PROC SQL;
   CREATE TABLE WORK.fof AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'FOF';
QUIT;
```

 # SPLIT DATA INTO THEMES: WDPROP (WOODY PROPORTION)

```
/*   START OF NODE: Query Builder51   */

PROC SQL;
   CREATE TABLE WORK.wdprop AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'WDPROP';
```

```
QUIT;
```

 # SPLIT DATA INTO THEMES: FO (FORESTRY)

```
/*   START OF NODE: Query Builder52   */

PROC SQL;
   CREATE TABLE WORK.fo AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'FO';
QUIT;
```

 # SPLIT DATA INTO THEMES: O (UNCATEGORIZED)

```
/*   START OF NODE: Query Builder53   */

PROC SQL;
   CREATE TABLE WORK.o AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'O';
QUIT;
```

 # SPLIT DATA INTO THEMES: HEATHER

```
/*   START OF NODE: Filter and Sort15   */

PROC SQL;
   CREATE TABLE WORK.heather AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
```

```
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'HEATHER';
QUIT;
```

 # SPLIT DATA INTO THEMES: LSC (LANDSCAPE)

```
/*   START OF NODE: Query Builder38   */

PROC SQL;
   CREATE TABLE WORK.All_POINTS_LSC AS
   SELECT DISTINCT t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.POINT_X,
          t1.POINT_Y
      FROM WORK.LC_DATA t1;
QUIT;
```

 # JOIN ALL THEMED TABLES INTO OUTPUT

```
/*   START OF NODE: Query Builder40   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_ALL_POINTS_LSC AS
   SELECT DISTINCT t1.HABT_GROUP,
          t8.SP_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.POINT_X,
          t1.POINT_Y,
          t8.CODE AS HABT,
          t8.DESCRIPTION AS HABT_DESC,
          t8.LUSE,
          t2.CODE AS USE,
          t2.DESCRIPTION AS USE_DESC,
          t2.Number AS NO_HORSES,
          t3.CODE AS CANOPY,
          t3.DESCRIPTION AS CANOPY_DESC,
          t4.CODE AS DESC,
          t4.DESCRIPTION AS DESC_DESC,
          t5.CODE AS FOF,
          t5.DESCRIPTION AS FOF_DESC,
          t6.CODE AS FORBS,
          t6.DESCRIPTION AS FORBS_DESC,
          t7.CODE AS HEATHER,
          t7.DESCRIPTION AS HEATHER_DESC,
          t10.CODE AS OTHER,
          t10.DESCRIPTION AS OTHER_DESC,
          t11.CODE AS PEAT,
          t11.DESCRIPTION AS PEAT_DESC,
          t12.CODE AS OTHER_USE,
          t12.DESCRIPTION AS OTHER_USE_DESC,
          t13.CODE AS WEEDPROP,
          t13.DESCRIPTION AS WEEDPROPR_DESC
      FROM WORK.ALL_POINTS_LSC t1
           LEFT JOIN WORK.AGUSE2 t2 ON (t1.HABT_GROUP = t2.HABT_GROUP)
           LEFT JOIN WORK.CAN t3 ON (t1.HABT_GROUP = t3.HABT_GROUP)
           LEFT JOIN WORK.DESC t4 ON (t1.HABT_GROUP = t4.HABT_GROUP)
```

```
            LEFT JOIN WORK.FOF t5 ON (t1.HABT_GROUP = t5.HABT_GROUP)
            LEFT JOIN WORK.FORBS t6 ON (t1.HABT_GROUP = t6.HABT_GROUP)
            LEFT JOIN WORK.HEATHER t7 ON (t1.HABT_GROUP = t7.HABT_GROUP)
            LEFT JOIN WORK.LC_HABTS t8 ON (t1.HABT_GROUP = t8.HABT_GROUP) AND
(t1.HABT_GROUP = t8.HABT_GROUP)
            LEFT JOIN WORK.MGT t9 ON (t1.HABT_GROUP = t9.HABT_GROUP)
            LEFT JOIN WORK.O t10 ON (t1.HABT_GROUP = t10.HABT_GROUP)
            LEFT JOIN WORK.USE t12 ON (t1.HABT_GROUP = t12.HABT_GROUP)
            LEFT JOIN WORK.WDPROP t13 ON (t1.HABT_GROUP = t13.HABT_GROUP)
            LEFT JOIN WORK.PEAT t11 ON (t1.HABT_GROUP = t11.HABT_GROUP);
QUIT;
```

KH5: Code to create Land Cover species attribute table

 # SELECT RECORDED SPECIES FROM MAIN DATA TABLE

```
/*   START OF NODE: Query Builder54   */

PROC SQL;
   CREATE TABLE WORK.sp AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'SP';
QUIT;
```

 # SELECT RECORDED SPECIES HEIGHTS FROM MAIN DATA TABLE

```
/*   START OF NODE: Query Builder55   */

PROC SQL;
   CREATE TABLE WORK.ht AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
```

```
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'HT';
QUIT;
```

 # SELECT RECORDED SPECIES AGE FROM MAIN DATA TABLE

```
/*   START OF NODE: Filter and Sort13   */

PROC SQL;
   CREATE TABLE WORK.age AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'AGE';
QUIT;
```

 # SELECT RECORDED SPECIES PROPORTION FROM MAIN DATA TABLE

```
/*   START OF NODE: Filter and Sort14   */

PROC SQL;
   CREATE TABLE WORK.prop AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'PROP';
QUIT;
```

 # SELECT RECORDED SPECIES % COVER FROM MAIN DATA TABLE

```
/*   START OF NODE: Filter and Sort16   */

PROC SQL;
   CREATE TABLE WORK.cover AS
```

```
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'COVER';
QUIT;
```

 # SELECT RECORDED SPECIES MANAGEMENT FROM MAIN DATA TABLE

```
/*   START OF NODE: Filter and Sort17   */

PROC SQL;
   CREATE TABLE WORK.FILTER_FOR_LC_DATA AS
   SELECT t1.Idno,
          t1.SP_GROUP,
          t1.HABT_GROUP,
          t1.SERIES_NUM,
          t1.CODE_TYPE,
          t1.GRIDCODE,
          t1.FAB_POSITION,
          t1.HAB_TYPE,
          t1.CODE,
          t1.DESCRIPTION,
          t1.ADD_CODE,
          t1.ADD_DESCRIPTION,
          t1.POINT_X,
          t1.POINT_Y,
          t1.CAT,
          t1.LUSE
      FROM WORK.LC_DATA t1
      WHERE t1.CAT = 'MGT';
QUIT;
```

 # CREATE LIST OF ALL EXPECTED RECORDED LOCATIONS

```
/*   START OF NODE: Query Builder43   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_LC_DATA AS
   SELECT DISTINCT t1.SERIES_NUM,
          t1.GRIDCODE,
          t1.POINT_X,
          t1.POINT_Y,
          t1.SP_GROUP
      FROM WORK.LC_DATA t1;
QUIT;
```

 # JOIN SPECIES DATA TO RECORDING LOCATIONS INFORMATION

```
/*   START OF NODE: Query Builder46   */
```

```
PROC SQL;
   CREATE TABLE WORK.SP2(label="QUERY_FOR_LC_DATA") AS
   SELECT DISTINCT t1.SP_GROUP AS SP_GROUP1,
          t1.SERIES_NUM,
          t1.GRIDCODE,
          t1.POINT_X,
          t1.POINT_Y,
          t2.CODE AS SPECIES,
          t2.DESCRIPTION
      FROM WORK.QUERY_FOR_LC_DATA t1
           LEFT JOIN WORK.SP t2 ON (t1.SERIES_NUM = t2.SERIES_NUM) AND (t1.GRIDCODE
= t2.GRIDCODE) AND (t1.SP_GROUP =
          t2.SP_GROUP);
QUIT;
```

 # ASSOCIATE HEIGHT, AGE, MANAGEMENT, COVER AND PROPORTION DATA TO SPECIES

```
/*   START OF NODE: Query Builder47   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_SP2_0000(label="QUERY_FOR_SP2") AS
   SELECT DISTINCT t1.SP_GROUP1,
          t1.SERIES_NUM,
          t1.GRIDCODE,
          t1.POINT_X,
          t1.POINT_Y,
          t1.SPECIES,
          t1.DESCRIPTION AS SPECIES_DESC,
          t2.CODE AS AGE,
          t2.DESCRIPTION AS AGE_DESC,
          t3.CODE AS COVER,
          t3.DESCRIPTION AS COVER_DESC,
          t4.CODE AS MGT,
          t4.DESCRIPTION AS MGT_DESC,
          t5.CODE AS HEIGHT,
          t5.DESCRIPTION AS HEIGHT_DESC
      FROM WORK.SP2 t1
           LEFT JOIN WORK.AGE t2 ON (t1.SP_GROUP1 = t2.SP_GROUP)
           LEFT JOIN WORK.COVER t3 ON (t1.SP_GROUP1 = t3.SP_GROUP)
           LEFT JOIN WORK.FILTER_FOR_LC_DATA t4 ON (t1.SP_GROUP1 = t4.SP_GROUP)
           LEFT JOIN WORK.HT t5 ON (t1.SP_GROUP1 = t5.SP_GROUP);
QUIT;
```

 # FILTER OUT ROWS WHERE NO DATA HAS BEEN RECORDED AND CREATE FINAL OUTPUT

```
/*   START OF NODE: Filter and Sort18   */

PROC SQL;
   CREATE TABLE WORK.FILTER_FOR_QUERY_FOR_SP2 AS
   SELECT t1.SP_GROUP1,
          t1.SERIES_NUM,
          t1.GRIDCODE,
          t1.SPECIES,
          t1.SPECIES_DESC,
          t1.AGE,
          t1.AGE_DESC,
          t1.COVER,
          t1.COVER_DESC,
          t1.MGT,
          t1.MGT_DESC,
          t1.HEIGHT,
          t1.HEIGHT_DESC
```

```
      FROM WORK.QUERY_FOR_SP2_0000 t1
      WHERE t1.SPECIES NOT IS MISSING OR t1.AGE NOT IS MISSING OR t1.COVER NOT IS
MISSING OR t1.MGT NOT IS MISSING OR
           t1.HEIGHT NOT IS MISSING;
QUIT;
```

KHS6: Key Habitats, code to transpose Boundary Codes

 # IMPORT DATA INTO SAS PROJECT: BOUNDARY RECORDED DATA

```
/*   START OF NODE: Import Data (BD_data.xlsx[final])   */

DATA WORK.BD_data;
    LENGTH
        Concat           $ 6
        Group2             8
        Group              8
        Sq               $ 5
        Grid             $ 1
        Fab                8
        Cat              $ 5
        Code               8
        Desc             $ 58
        LUSE             $ 3 ;
    FORMAT
        Concat           $CHAR6.
        Group2           BEST12.
        Group            BEST12.
        Sq               $CHAR5.
        Grid             $CHAR1.
        Fab              BEST12.
        Cat              $CHAR5.
        Code             BEST12.
        Desc             $CHAR58.
        LUSE             $CHAR3. ;
    INFORMAT
        Concat           $CHAR6.
        Group2           BEST12.
        Group            BEST12.
        Sq               $CHAR5.
        Grid             $CHAR1.
        Fab              BEST12.
        Cat              $CHAR5.
        Code             BEST12.
        Desc             $CHAR58.
        LUSE             $CHAR3. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\BD_data-
0500d165c31440248ee77f4a60a343bc.txt'
        LRECL=90
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Concat           : $CHAR6.
        Group2           : BEST32.
        Group            : BEST32.
        Sq               : $CHAR5.
        Grid             : $CHAR1.
        Fab              : BEST32.
        Cat              : $CHAR5.
        Code             : BEST32.
        Desc             : $CHAR58.
        LUSE             : $CHAR3. ;
RUN;
```

 # IMPORT BOUNDARY LAND USE CATEGORY LOOKUP INTO SAS PROJECT

```
/*   START OF NODE: Import Data (BD_Code_list.xlsx[Columns])   */

DATA WORK.BD_Code_list_columns;
    LENGTH
        Columns          $ 5
        F2               $ 20 ;
    FORMAT
        Columns          $CHAR5.
        F2               $CHAR20. ;
    INFORMAT
        Columns          $CHAR5.
        F2               $CHAR20. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\BD_Code_list-
6637da37d44d41579b6678831313444b.txt'
        LRECL=9
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Columns          : $CHAR5.
        F2               : $CHAR20. ;
RUN;
```

 # IMPORT DATA INTO SAS PROJECT: BOUNDARY DESCRIPTION CODE LIST

```
/*   START OF NODE: Import Data (BD_Code_list.xlsx[BD_codes])   */

DATA WORK.BD_Code_list;
    LENGTH
        Code               8
        Desc             $ 58
        Cat              $ 5
        LUSE             $ 3 ;
    FORMAT
        Code             BEST12.
        Desc             $CHAR58.
        Cat              $CHAR5.
        LUSE             $CHAR3. ;
    INFORMAT
        Code             BEST12.
        Desc             $CHAR58.
        Cat              $CHAR5.
        LUSE             $CHAR3. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\BD_Code_list-
20b23ab254c94d1d9d9eb29fe0ec3659.txt'
        LRECL=70
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Code             : BEST32.
        Desc             : $CHAR58.
        Cat              : $CHAR5.
        LUSE             : $CHAR3. ;
RUN;
```

 # IMPORT DATA INTO SAS PROJECT: VEGETATION SPECIES WHERE NOT BRACKEN

```
/*   START OF NODE: Import Data (BD_data.xlsx[species not bracken])   */

DATA WORK.sp_no_brac;
```

```
    LENGTH
        Concat           $ 6
        Group3             8
        Sp_group           8
        Group2             8
        Group              8
        Sq               $ 5
        Grid             $ 1
        Fab                8
        Cat              $ 4
        Code               8
        Desc             $ 54
        LUSE             $ 1 ;
    FORMAT
        Concat           $CHAR6.
        Group3           BEST12.
        Sp_group         BEST12.
        Group2           BEST12.
        Group            BEST12.
        Sq               $CHAR5.
        Grid             $CHAR1.
        Fab              BEST12.
        Cat              $CHAR4.
        Code             BEST12.
        Desc             $CHAR54.
        LUSE             $CHAR1. ;
    INFORMAT
        Concat           $CHAR6.
        Group3           BEST12.
        Sp_group         BEST12.
        Group2           BEST12.
        Group            BEST12.
        Sq               $CHAR5.
        Grid             $CHAR1.
        Fab              BEST12.
        Cat              $CHAR4.
        Code             BEST12.
        Desc             $CHAR54.
        LUSE             $CHAR1. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\BD_data-
cd36d10637ec4f189306edf218e21833.txt'
        LRECL=89
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Concat           : $CHAR6.
        Group3           : BEST32.
        Sp_group         : BEST32.
        Group2           : BEST32.
        Group            : BEST32.
        Sq               : $CHAR5.
        Grid             : $CHAR1.
        Fab              : BEST32.
        Cat              : $CHAR4.
        Code             : BEST32.
        Desc             : $CHAR54.
        LUSE             : $CHAR1. ;
RUN;
```

 # IMPORT DATA INTO SAS PROJECT: BRACKEN DATA

```
/*   START OF NODE: Import Data (BD_data.xlsx[bracken])   */

DATA WORK.bracken_hts;
    LENGTH
```

```
        'Bracken ht'n    $ 9
        'Bracken ht code'n   8
        Concat           $ 4
        Group2             8
        Group              8
        Sq               $ 3
        Grid             $ 1
        Fab                8
        Cat              $ 4
        Code               8
        Desc             $ 15
        LUSE             $ 1 ;
    FORMAT
        'Bracken ht'n    $CHAR9.
        'Bracken ht code'n BEST12.
        Concat           $CHAR4.
        Group2           BEST12.
        Group            BEST12.
        Sq               $CHAR3.
        Grid             $CHAR1.
        Fab              BEST12.
        Cat              $CHAR4.
        Code             BEST12.
        Desc             $CHAR15.
        LUSE             $CHAR1. ;
    INFORMAT
        'Bracken ht'n    $CHAR9.
        'Bracken ht code'n BEST12.
        Concat           $CHAR4.
        Group2           BEST12.
        Group            BEST12.
        Sq               $CHAR3.
        Grid             $CHAR1.
        Fab              BEST12.
        Cat              $CHAR4.
        Code             BEST12.
        Desc             $CHAR15.
        LUSE             $CHAR1. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\BD_data-
342c1244b5294f26a8200d2012b3474c.txt'
        LRECL=60
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        'Bracken ht'n    : $CHAR9.
        'Bracken ht code'n : BEST32.
        Concat           : $CHAR4.
        Group2           : BEST32.
        Group            : BEST32.
        Sq               : $CHAR3.
        Grid             : $CHAR1.
        Fab              : BEST32.
        Cat              : $CHAR4.
        Code             : BEST32.
        Desc             : $CHAR15.
        LUSE             : $CHAR1. ;
RUN;
```

 # IMPORT DATA INTO SAS PROJECT: HTS (HEIGHTS)

```
/*   START OF NODE: Import Data (BD_data.xlsx[hts])   */

DATA WORK.HTS;
    LENGTH
        Concat           $ 6
```

```
        Group2            8
        Group             8
        Sq              $ 5
        Grid            $ 1
        Fab               8
        Cat             $ 2
        Code              8
        Desc            $ 9
        LUSE            $ 1 ;
    FORMAT
        Concat          $CHAR6.
        Group2          BEST12.
        Group           BEST12.
        Sq              $CHAR5.
        Grid            $CHAR1.
        Fab             BEST12.
        Cat             $CHAR2.
        Code            BEST12.
        Desc            $CHAR9.
        LUSE            $CHAR1. ;
    INFORMAT
        Concat          $CHAR6.
        Group2          BEST12.
        Group           BEST12.
        Sq              $CHAR5.
        Grid            $CHAR1.
        Fab             BEST12.
        Cat             $CHAR2.
        Code            BEST12.
        Desc            $CHAR9.
        LUSE            $CHAR1. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\BD_data-
5914032239804dafba9b997859f658ea.txt'
        LRECL=42
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Concat          : $CHAR6.
        Group2          : BEST32.
        Group           : BEST32.
        Sq              : $CHAR5.
        Grid            : $CHAR1.
        Fab             : BEST32.
        Cat             : $CHAR2.
        Code            : BEST32.
        Desc            : $CHAR9.
        LUSE            : $CHAR1. ;
RUN;
```

 # IMPORT DATA INTO SAS PROJECT: CANOPIES

```
/*   START OF NODE: Import Data (BD_data.xlsx[canopies])   */

DATA WORK.cans;
    LENGTH
        Concat          $ 4
        Group2            8
        Group             8
        Sq              $ 3
        Grid            $ 1
        Fab               8
        Cat             $ 3
        Code              8
        Desc            $ 21
        LUSE            $ 1 ;
```

```
    FORMAT
        Concat           $CHAR4.
        Group2           BEST12.
        Group            BEST12.
        Sq               $CHAR3.
        Grid             $CHAR1.
        Fab              BEST12.
        Cat              $CHAR3.
        Code             BEST12.
        Desc             $CHAR21.
        LUSE             $CHAR1. ;
    INFORMAT
        Concat           $CHAR4.
        Group2           BEST12.
        Group            BEST12.
        Sq               $CHAR3.
        Grid             $CHAR1.
        Fab              BEST12.
        Cat              $CHAR3.
        Code             BEST12.
        Desc             $CHAR21.
        LUSE             $CHAR1. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG8444\BD_data-
98357353dc164b23b79fdd1b7bf1dfa5.txt'
        LRECL=51
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        Concat           : $CHAR4.
        Group2           : BEST32.
        Group            : BEST32.
        Sq               : $CHAR3.
        Grid             : $CHAR1.
        Fab              : BEST32.
        Cat              : $CHAR3.
        Code             : BEST32.
        Desc             : $CHAR21.
        LUSE             : $CHAR1. ;
RUN;
```

 # REMOVE DUPLICATES FROM LAND COVER DESCRIPTION CODE LIST

```
/*   START OF NODE: Query Builder18   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_FILTER_FOR_BD_CODES_WI(label="QUERY_FOR_FILTER_FOR_BD_CODES_WITH_PCO
DE") AS
   SELECT DISTINCT t1.CODE,
          t1.DESCRIPTION
      FROM WORK.FILTER_FOR_BD_CODES_WITH_PCODE t1
      ORDER BY t1.DESCRIPTION;
QUIT;
```

 # JOIN LAND USE CATEGORY TO RECORDED BOUNDARY DATA

```
/*   START OF NODE: Query Builder19   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_FILTER_FOR_BD_COD_0000(label="QUERY_FOR_FILTER_FOR_BD_CODES_WITH_PCO
DE") AS
   SELECT t1.SERIES_NUM,
          t1.GRIDCODE,
          t1.FAB_POSITION,
```

```
          t2.Cat,
          t1.CODE,
          t1.DESCRIPTION,
          t2.LUSE
      FROM WORK.FILTER_FOR_BD_CODES_WITH_PCODE t1, WORK.BD_CODE_LIST t2
      WHERE (t1.DESCRIPTION = t2.Desc AND t1.CODE = t2.Code)
      ORDER BY t1.SERIES_NUM,
               t1.GRIDCODE,
               t1.FAB_POSITION;
QUIT;
```

 # SPLIT DATA INTO THEMES: HABT (HABITATS)

```
/*   START OF NODE: Query Builder21   */

PROC SQL;
   CREATE TABLE WORK.bdhabts AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.BD_DATA t1
      WHERE t1.Cat = 'HABT';
QUIT;
```

 # SPLIT DATA INTO THEMES: STOCK

```
/*   START OF NODE: Query Builder22   */

PROC SQL;
   CREATE TABLE WORK.bdstock(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.BD_DATA t1
      WHERE t1.Cat = 'STOCK';
QUIT;
```

 # SPLIT DATA INTO THEMES: USE

```
/*   START OF NODE: Query Builder23   */

PROC SQL;
   CREATE TABLE WORK.bduse(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
```

```
      FROM WORK.BD_DATA t1
      WHERE t1.Cat = 'USE';
QUIT;
```

 # SPLIT DATA INTO THEMES: DESC

```
/*   START OF NODE: Query Builder24   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_BD_DATA_0002(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.BD_DATA t1
      WHERE t1.Cat = 'DESC';
QUIT;
```

 # SPLIT DATA INTO THEMES: BDMGT (BOUNDARY MANAGEMENT)

```
/*   START OF NODE: Query Builder25   */

PROC SQL;
   CREATE TABLE WORK.bdmgt(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.BD_DATA t1
      WHERE t1.Cat = 'MGT';
QUIT;
```

 # SPLIT DATA INTO THEMES: O (UNCATEGORISED)

```
/*   START OF NODE: Query Builder26   */

PROC SQL;
   CREATE TABLE WORK.bdo(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.BD_DATA t1
      WHERE t1.Cat = 'O';
QUIT;
```

 # SPLIT DATA INTO THEMES: GAPS

```
/*   START OF NODE: Query Builder28   */
```

```
PROC SQL;
   CREATE TABLE WORK.bdgaps(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.BD_DATA t1
      WHERE t1.Cat = 'GAPS';
QUIT;

/*   START OF NODE: Query Builder13   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_BD_DATA_0009(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.BD_DATA t1
      WHERE t1.Desc = 'Mixed hedge';
QUIT;

/*   START OF NODE: Query Builder50   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_BD_DATA AS
   SELECT DISTINCT t1.Cat,
          t1.Desc
      FROM WORK.BD_DATA t1;
QUIT;
```

 # CREATE FINAL BOUNDARY FEATURE TABLE BY JOINING DATA ASSEMBLED IN TABLES ABOVE

```
/*   START OF NODE: Query Builder31   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_BDHABTS AS
   SELECT DISTINCT t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Code AS HABT,
          t1.Desc AS HABT_DESC,
          t1.LUSE,
          t2.Code AS BRACKEN,
          t2.Desc AS BRACKEN_DESC,
          t9.'Bracken ht'n AS BRACKEN_HT,
          t9.'Bracken ht code'n AS BRACKEN_HT_CODE,
          t3.Code AS GAPS,
          t3.Desc AS GAPS_DESC,
          t4.Code AS MGT,
          t4.Desc AS MGT_DESC,
          t6.Code AS STOCK,
          t6.Desc AS STOCK_DESC,
          t7.Code AS USE,
```

```
          t7.Desc AS USE_DESC,
          t8.Desc1 AS DESC1,
          t8.Desc_code1 AS DESC1_DESC,
          t8.Desc2 AS DESC2,
          t8.Desc_code2 AS DESC2_DESC,
          t8.Desc_code3 AS DESC3,
          t8.Desc3 AS DESC3_DESC,
          t5.Desc AS OTHER_DESC,
          t5.Code AS OTHER,
          t10.Code AS CANOPY,
          t10.Desc AS CANOPY_DESC,
          t11.Code AS HEIGHT,
          t11.Desc AS HEIGHT_DESC
      FROM WORK.BDHABTS t1
           LEFT JOIN WORK.BDBRACKEN t2 ON (t1.Sq = t2.Sq) AND (t1.Grid = t2.Grid)
           LEFT JOIN WORK.BDGAPS t3 ON (t1.Sq = t3.Sq) AND (t1.Grid = t3.Grid)
           LEFT JOIN WORK.BDMGT t4 ON (t1.Sq = t4.Sq) AND (t1.Grid = t4.Grid)
           LEFT JOIN WORK.BDO t5 ON (t1.Sq = t5.Sq) AND (t1.Grid = t5.Grid)
           LEFT JOIN WORK.BDSTOCK t6 ON (t1.Sq = t6.Sq) AND (t1.Grid = t6.Grid) AND
(t1.Sq = t6.Sq)
           LEFT JOIN WORK.BDUSE t7 ON (t1.Grid = t7.Grid) AND (t1.Sq = t7.Sq)
           LEFT JOIN WORK.BRACKEN_HTS t9 ON (t1.Sq = t9.Sq) AND (t1.Grid = t9.Grid)
           LEFT JOIN WORK.DESCS t8 ON (t1.Grid = t8.Grid) AND (t1.Sq = t8.Sq)
           LEFT JOIN WORK.CANS t10 ON (t1.Sq = t10.Sq) AND (t1.Grid = t10.Grid)
           LEFT JOIN WORK.HTS t11 ON (t1.Sq = t11.Sq) AND (t1.Grid = t11.Grid)
      ORDER BY t1.Sq,
               t1.Grid;
QUIT;
```

 # APPEND LOCATION DATA TO FINAL BOUNDARY FEATURE TABLE

```
/*   START OF NODE: Query Builder32   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_BDHABTS_0001(label="QUERY_FOR_BDHABTS") AS
   SELECT DISTINCT t1.Sq,
          t1.Group2,
          t1.Grid,
          t1.HABT,
          t1.HABT_DESC,
          t1.LUSE,
          t1.BRACKEN,
          t1.BRACKEN_DESC,
          t1.BRACKEN_HT,
          t1.BRACKEN_HT_CODE,
          t1.GAPS,
          t1.GAPS_DESC,
          t1.MGT,
          t1.MGT_DESC,
          t1.STOCK,
          t1.STOCK_DESC,
          t1.USE,
          t1.USE_DESC,
          t1.DESC1,
          t1.DESC1_DESC,
          t1.DESC2,
          t1.DESC2_DESC,
          t1.DESC3,
          t1.DESC3_DESC,
          t1.OTHER_DESC,
          t1.OTHER,
          t1.CANOPY,
          t1.CANOPY_DESC,
          t1.HEIGHT,
          t1.HEIGHT_DESC,
          t2.POINT_X,
          t2.POINT_Y
      FROM WORK.QUERY_FOR_BDHABTS t1
```

```
           LEFT JOIN WORK.BD_XY t2 ON (t1.Sq = t2.SERIES_NUM) AND (t1.Grid =
t2.GRIDCODE);
QUIT;

/*   START OF NODE: Query Builder56   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_BDHABTS_0002(label="QUERY_FOR_BDHABTS") AS
   SELECT DISTINCT t1.Sq AS SERIES_NUM,
          t1.Group2 AS SP_LINK,
          t1.Grid AS Gridcode,
          t1.HABT,
          t1.HABT_DESC,
          t1.LUSE,
          t1.BRACKEN,
          t1.BRACKEN_DESC,
          t1.BRACKEN_HT,
          t1.BRACKEN_HT_CODE,
          t1.GAPS,
          t1.GAPS_DESC,
          t1.MGT,
          t1.MGT_DESC,
          t1.STOCK,
          t1.STOCK_DESC,
          t1.USE,
          t1.USE_DESC,
          t1.DESC1,
          t1.DESC1_DESC,
          t1.DESC2,
          t1.DESC2_DESC,
          t1.DESC3,
          t1.DESC3_DESC,
          t1.OTHER_DESC,
          t1.OTHER,
          t1.CANOPY,
          t1.CANOPY_DESC,
          t1.POINT_X,
          t1.POINT_Y
      FROM WORK.QUERY_FOR_BDHABTS_0001 t1;
QUIT;

/*   START OF NODE: Query Builder59   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_BDHABTS_0004(label="QUERY_FOR_BDHABTS") AS
   SELECT DISTINCT t1.SERIES_NUM,
          t1.SP_LINK,
          t1.Gridcode,
          t1.HABT,
          t1.HABT_DESC,
          t1.LUSE,
          t1.BRACKEN,
          t1.BRACKEN_DESC,
          t1.BRACKEN_HT,
          t1.BRACKEN_HT_CODE,
          t1.GAPS,
          t1.GAPS_DESC,
          t1.MGT,
          t1.MGT_DESC,
          t1.STOCK,
          t1.STOCK_DESC,
          t1.USE,
          t1.USE_DESC,
          t1.DESC1,
          t1.DESC1_DESC,
          t1.DESC2,
          t1.DESC2_DESC,
          t1.DESC3,
```

```
          t1.DESC3_DESC,
          t1.OTHER_DESC,
          t1.OTHER,
          t1.CANOPY,
          t1.CANOPY_DESC,
          t1.POINT_X,
          t1.POINT_Y,
          t2.Code,
          t2.Desc
      FROM WORK.QUERY_FOR_BDHABTS_0002 t1
           LEFT JOIN WORK.HTS t2 ON (t1.SP_LINK = t2.Group2);
QUIT;
```

KHS7: Transpose primary codes: Boundary Species

 # SPLIT DATA INTO THEMES: SELECT BRACKEN DATA (AS IS AN ATTRIBUTE OF ANOTHER FEATURE)

```
/*   START OF NODE: Query Builder29   */

PROC SQL;
   CREATE TABLE WORK.bdbracken(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.QUERY_FOR_BD_DATA_0002 t1
      WHERE t1.Desc IN
           (
           'Bracken - scattered',
           'Bracken present'
           );
QUIT;
```

 # SPLIT DATA INTO THEMES: SELECT DATA WHERE NOT BRACKEN

```
/*   START OF NODE: Query Builder30   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_BD_DATA_0008(label="QUERY_FOR_BD_DATA") AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.QUERY_FOR_BD_DATA_0002 t1
      WHERE t1.Desc NOT IN
           (
           'Bracken - scattered',
           'Bracken present'
           );
QUIT;
```

 # TRANSPOSE SPECIES DATA DESCRIPTIONS

```
/*   START OF NODE: Transpose28   */
```

```
PROC SORT
       DATA=WORK.QUERY_FOR_BD_DATA_0008(KEEP=Desc Group2 Sq Grid)
       OUT=WORK.SORTTempTableSorted
       ;
       BY Group2 Sq Grid;
RUN;
PROC TRANSPOSE DATA=WORK.SORTTempTableSorted
       OUT=WORK.DESC_DESCS(LABEL="Transposed WORK.QUERY_FOR_BD_DATA_0008")
       PREFIX=Column
       NAME=Source
       LABEL=Label
;
       BY Group2 Sq Grid;
       VAR Desc;

RUN; QUIT;
```

 # TRANSPOSE SPECIES DATA CODES

```
/*   START OF NODE: Transpose29   */

PROC SORT
       DATA=WORK.QUERY_FOR_BD_DATA_0008(KEEP=Code Group2 Grid Sq)
       OUT=WORK.SORTTempTableSorted
       ;
       BY Group2 Grid Sq;
RUN;
PROC TRANSPOSE DATA=WORK.SORTTempTableSorted
       OUT=WORK.DESC_CODES(LABEL="Transposed WORK.QUERY_FOR_BD_DATA_0008")
       PREFIX=Column
       NAME=Source
       LABEL=Label
;
       BY Group2 Grid Sq;
       VAR Code;

RUN; QUIT;
```

 # JOIN TRANSPOSED SPECIES DESCRIPTIONS AND CODES

```
/*   START OF NODE: Query Builder27   */

PROC SQL;
   CREATE TABLE WORK.descs AS
   SELECT t1.Sq,
          t1.Grid,
          t1.Group2,
          t1.Source,
          t2.Column1 AS Desc1,
          t1.Column1 AS Desc_code1,
          t1.Column2 AS Desc2,
          t2.Column2 AS Desc_code2,
          t1.Column3 AS Desc3,
          t2.Column3 AS Desc_code3
      FROM WORK.DESC_DESCS t1
           INNER JOIN WORK.DESC_CODES t2 ON (t1.Group2 = t2.Group2);
QUIT;
```

 # CREATE FINAL SPECIES TABLE BY FILTERING SPECIES INFORMATION

```
/*   START OF NODE: Query Builder14   */

PROC SQL;
   CREATE TABLE WORK.SP AS
   SELECT t1.Concat,
          t1.Group3,
```

```
          t1.Group2,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.SP_NO_BRAC t1
      WHERE t1.Cat = 'SP';
QUIT;
```

 # FILTER AGE DATA INTO SEPARATE TABLE

```
/*   START OF NODE: Query Builder15   */

PROC SQL;
   CREATE TABLE WORK.AGE AS
   SELECT t1.Concat,
          t1.Group3,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.SP_NO_BRAC t1
      WHERE t1.Cat = 'AGE';
QUIT;
```

 # FILTER PROPORTION DATA INTO SEPARATE TABLE

```
/*   START OF NODE: Query Builder16   */

PROC SQL;
   CREATE TABLE WORK.PROP AS
   SELECT t1.Concat,
          t1.Group2,
          t1.Group3,
          t1.Group,
          t1.Sq,
          t1.Grid,
          t1.Fab,
          t1.Cat,
          t1.Code,
          t1.Desc,
          t1.LUSE
      FROM WORK.SP_NO_BRAC t1
      WHERE t1.Cat = 'PROP';
QUIT;
```

#JOIN AGE AND PROPORTION DATA & DESCRIPTIONS TO FINAL TABLE

```
/*   START OF NODE: Query Builder17   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_SP AS
   SELECT DISTINCT t1.Sq,
          t1.Grid,
          t1.Group2,
          t1.Group3,
          t1.Code AS SPECIES,
          t1.Desc AS SPECIES_DESC,
          t2.Code AS AGE,
          t2.Desc AS AGE_DESC,
```

```
          t3.Code AS PROPORTION,
          t3.Desc AS PROP_DESC
      FROM WORK.SP t1
           LEFT JOIN WORK.AGE t2 ON (t1.Group3 = t2.Group3)
           LEFT JOIN WORK.PROP t3 ON (t1.Group3 = t3.Group3);
QUIT;
```

KHS8: Code to create perform updates on 'Key Habitat' survey Land Use codes

/********Script to perform updates on Key Habs LUSE codes C.Wood 11/7/13 ****/

 # IMPORT LOOKUP TABLE DATA
/* Landscape codes*/

```
/*Bunce surveys*/ LIBNAME BUNCE ORACLE PATH='fegen'  SCHEMA='bunce_surveys'
USER=xxxxx'  PASSWORD='xxxx';
PROC IMPORT OUT= WORK.LCLUSE
            DATAFILE= "N:\Documents\Current_projects\ILO
work\CEH_Gateway\Ingestion\Key_Habitats\Final_data\FINAL_KH_FINAL\Transposing_Codes
\FINAL\LC_LUSE_CODES.csv"
            DBMS=CSV REPLACE;
     GETNAMES=YES;
     DATAROW=2;
RUN;
```

 # UPDATE LAND USE CODES IN MAIN DATA TABLE TO ENSURE CONSISTENT CODES AND POPULATE WHERE MISSING
 # ARABLE CODES

```
proc sql;
    update BUNCE.Key_habitat_lc_features a
    set luse = (select LUSE from WORK.LCLUSE b
                          where a.HABT_DESC=b.HABT_DESC)
    where LUSE ="AC";
     quit;
```

 # AGRICULTURE/NATURAL VEG CODES (ORCHARD)

```
proc sql;
    update bunce.Key_habitat_lc_features
    set LUSE = "AN"
    where USE_DESC = "Orchard" and LUSE is null;
quit;
```

 # AGRICULTURE USE CODES (SET-ASIDE)

```
proc sql;
    update bunce.Key_habitat_lc_features
    set LUSE = "AGU"
    where USE_DESC = "Set aside" and LUSE is null;
quit;
```

 # AGRICULTURE/NATURAL VEG CODES (COMMERCIAL HORTICULTURE)

```
proc sql;
    update bunce.Key_habitat_lc_features
    set LUSE = "AN"
    where USE_DESC = "Commercial horticulture" and LUSE is null;
quit;
```

#FORESTRY IN SPECIFIC SQUARES

```
proc sql;
    update bunce.Key_habitat_lc_features
    set LUSE = "FOF"
    where SERIES_NUM IN ("BD15","MN4","TN3","TN4")and LUSE is null;
quit;
```

#FORESTRY IN SPECIFIC SQUARES

```
proc sql;
    update bunce.Key_habitat_lc_features
    set LUSE = "FO"
    where SERIES_NUM = "MN6" and LUSE is null;
quit;
```

#STRUCTURES

```
proc sql;
    update bunce.Key_habitat_lc_features
    set LUSE = "ST"
    where USE in (430,431)and LUSE is null;
quit;
```

#UNSURVEYED

```
proc sql;
    update bunce.Key_habitat_lc_features
    set LUSE = "US"
    where LUSE is null;
quit;
```

/*Boundary codes*/

 # SET SEA WALL to WALL NOT FENCE

```
proc sql;
    update bunce.Key_habitat_bd_features
    set LUSE = "W"
    where HABT_DESC="Sea wall" and LUSE ="F";
quit;
```

 # SET HEDGE TO HEDGE NOT WALL

```
proc sql;
    update bunce.Key_habitat_bd_features
    set LUSE = "WLF"
    where HABT_DESC="Hedge >2m wide" and LUSE ="W";
quit;
```

 # IMPORT LOOKUP TABLE DATA

```
/* PROC IMPORT OUT= WORK.LC_SP
            DATAFILE= "N:\Documents\Current_projects\ILO work\CEH_Gatewa
y\Ingestion\Key_Habitats\Final_data\FINAL_KH_FINAL\Transposing_Codes\FINAL\SP_LINK_
codes.xls"
            DBMS=EXCEL REPLACE;
     SHEET="LC_sp$";
     GETNAMES=YES;
     MIXED=NO;
     SCANTEXT=YES;
     USEDATE=YES;
     SCANTIME=YES;
RUN;

PROC IMPORT OUT= WORK.BD_FEAT

DATAFILE= "N:\Documents\Current_projects\ILO work\CEH_Gatewa
y\Ingestion\Key_Habitats\Final_data\FINAL_KH_FINAL\Transposing_Codes\FINAL\SP_LINK_
codes.xls"
            DBMS=EXCEL REPLACE;
     SHEET="BD_feat$";
     GETNAMES=YES;
     MIXED=NO;
     SCANTEXT=YES;
     USEDATE=YES;
     SCANTIME=YES;
RUN;

PROC IMPORT OUT= WORK.BD_Sp
```

```
            DATAFILE= "N:\Documents\Current_projects\ILO 
work\CEH_Gateway\Ingestion\Key_Habitats\Final_data\FINAL_KH_FINAL\Transposing_Codes
\FINAL\SP_LINK_codes.xls"
            DBMS=EXCEL REPLACE;
     SHEET="BD_sp$";
     GETNAMES=YES;
     MIXED=NO;
     SCANTEXT=YES;
     USEDATE=YES;
     SCANTIME=YES;
RUN;
```

 # UPDATE SPECIES LINK FIELD TO LINK TABLES

```
proc sql;
    update BUNCE.Key_habitat_bd_features a
    set SP_LINK = (select Group2 from WORK.bd_feat b
                          where a.Series_num=b.Series_num and 
a.Gridcode=b.Gridcode)
    where SP_LINK is null;
     quit;

proc sql;
    update BUNCE.Key_habitat_bd_species a
    set SP_LINK = (select Group2 from WORK.bd_sp b
                          where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode 
and a.Species=b.Sp_no)
    where SP_LINK is null;
     quit;


 proc sql;
    update BUNCE.Key_habitat_lc_species a
    set SP_LINK = (select Group2 from WORK.lc_sp b
                          where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode 
and a.Species=b.Sp_no)
    where SP_LINK is null;
     quit;
        */
```

/* Update codes in new tables */

```
proc sql;
create table lcsp1 as
select distinct series_num, gridcode, species,species_desc from 
BUNCE.Key_habitat_lc_species
where species_desc is not null;
quit;

proc sql;
create table lcsp2 as
select distinct series_num, gridcode, species, mgt_desc from 
BUNCE.Key_habitat_lc_species
where species_desc is not null;
quit;

proc sql;
    update BUNCE.Key_habitat_lc_species_new a
    set Species_desc = (select Species_desc from lcsp1 b
                          where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode 
and a.Species=b.Species)
    where series_num not in("AN8","MD7","MD1");
     quit;

proc sql;
create table bdsp as
select distinct series_num, gridcode, species, species_desc from
```

```
BUNCE.Key_habitat_bd_species
where species_desc is not null;
quit;

proc sql;
    update BUNCE.Key_habitat_bd_species_new a
    set Species_desc = (select Species_desc from bdsp b
                          where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode
and a.Species=b.Species)
    where species_desc is not null;
     quit;

proc sql;
create table lchab as
select distinct series_num, gridcode, habt, habt_desc from
BUNCE.Key_habitat_lc_features
where habt_desc is not null;
quit;

proc sql;
    update BUNCE.Key_habitat_lc_features2 a
    set Habt_desc = (select habt_desc from lchab b
  where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.Habt=b.Habt);
     quit;

 proc sql;
create table lcuse as
select distinct habt_desc,luse  from
BUNCE.Key_habitat_lc_features
where habt_desc is not null;
quit;

proc sql;
update BUNCE.Key_habitat_lc_features2 a
 set LUSE = (select LUSE from lcuse b
 where a.Habt_desc=b.Habt_desc);
  quit;

proc sql;
create table desc2 as
select distinct series_num, gridcode, habt, desc_desc  from
BUNCE.Key_habitat_lc_features
where desc_desc is not null;
quit;

 proc sql;
    update BUNCE.Key_habitat_lc_features2 a
    set desc_desc = (select desc_desc from desc2 b
    where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.Habt=b.Habt);
     quit;

proc sql;
update BUNCE.Key_habitat_lc_features2
set FOF_DESC = "Glade"
 where FOF_DESC= "GLADE";
 quit;

proc sql;
create table ouse as
select distinct series_num, gridcode, habt, other_use, other_use_desc  from
BUNCE.Key_habitat_lc_features
where other_use_desc is not null;
quit;


proc sql;
    update BUNCE.Key_habitat_lc_features2 a
    set other_use_desc = (select other_use_desc from ouse b
```

```
    where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.Habt=b.Habt and
a.other_use=b.other_use);
     quit;

PROC SQL;
 CREATE TABLE WORK.crops AS SELECT distinct KEY_HABITAT_LC_FEATURES.SERIES_NUM,
        KEY_HABITAT_LC_FEATURES.GRIDCODE,
        KEY_HABITAT_LC_FEATURES.HABT,
        KEY_HABITAT_LC_FEATURES.HABT_DESC,
        KEY_HABITAT_LC_FEATURES.LUSE
 FROM BUNCE.KEY_HABITAT_LC_FEATURES AS KEY_HABITAT_LC_FEATURES
 WHERE KEY_HABITAT_LC_FEATURES.LUSE = "AC";
QUIT;

proc sql;
    update BUNCE.Key_habitat_lc_features2 a
    set HABT_DESC = (select Habt_desc from crops b
    where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode)
where a.habt_desc is null;
     quit;


proc sql;
    update BUNCE.Key_habitat_lc_features2 a
    set LUSE = (select LUSE from crops b
    where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode)
where LUSE is null;
     quit;


proc sql;
create table usedesc as
select distinct series_num, gridcode, habt, use, use_desc  from
BUNCE.Key_habitat_lc_features
where use_desc is not null;
quit;

proc sql;
    update BUNCE.Key_habitat_lc_features2 a
    set use_desc = (select use_desc from usedesc b
    where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.use=b.use);
     quit;

proc sql;
create table habs as
select distinct series_num, hab_type  from
BUNCE.Key_habitat_lc_features;
quit;

proc sql;
update BUNCE.Key_habitat_lc_features2 a
set hab_type = (select hab_type from habs b
where a.Series_num=b.Series_num);
 quit;

proc sql;
 update BUNCE.Key_habitat_bd_features2 a
set hab_type = (select hab_type from habs b
where a.Series_num=b.Series_num);
 quit;


proc sql;
create table bdhabs as
select distinct series_num,gridcode, habt, habt_desc  from
BUNCE.Key_habitat_bd_features_old;
quit;
```

```
proc sql;
 update BUNCE.Key_habitat_bd_features2 a
set habt_desc = (select habt_desc from bdhabs b
where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.Habt=b.Habt);
quit;


proc sql;
create table bdluse as
select distinct series_num,gridcode, habt, LUSE  from
BUNCE.Key_habitat_bd_features_old;
quit;

proc sql;
    update BUNCE.Key_habitat_bd_features2 a
    set luse = (select luse from bdluse b
    where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.Habt=b.Habt);
     quit;

 proc sql;
create table bdmgt as
select distinct series_num,gridcode, mgt, mgt_desc  from
BUNCE.Key_habitat_bd_features_old;
quit;

proc sql;
update BUNCE.Key_habitat_bd_features2 a
  set mgt_desc = (select mgt_desc from bdmgt b
  where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.Mgt=b.Mgt);
  quit;

proc sql;
update BUNCE.Key_habitat_bd_features2
set Use_desc = "Amenity"
 where Use_desc="AMENITY";
  quit;

proc sql;
create table bduse1 as
select distinct series_num,gridcode, habt,desc1_desc   from
BUNCE.Key_habitat_bd_features_old
where desc1_desc is not null
order by series_num, gridcode, habt;
quit;

proc sql;
create table bduse2 as
select distinct series_num,gridcode, habt,desc2_desc  from
BUNCE.Key_habitat_bd_features_old;
quit;

proc sql;
create table bduse3 as
select distinct series_num, gridcode, habt,desc3_desc  from
BUNCE.Key_habitat_bd_features_old
order by series_num,gridcode,habt;
quit;

proc sql;
update BUNCE.Key_habitat_bd_features2 a
set desc2_desc = (select desc2_desc from bduse2 b
 where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.habt=b.habt);
 quit;

proc sql;
update BUNCE.Key_habitat_bd_features2 a
 set desc3_desc = (select desc3_desc from bduse3 b
where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.habt=b.habt);
```

```
 quit;

proc sql;
create table bdouse as
select distinct series_num,gridcode, habt , other_desc  from
BUNCE.Key_habitat_bd_features_old;
quit;

proc sql;
update BUNCE.Key_habitat_bd_features2 a
 set other_desc = (select other_desc from bdouse b
 where a.Series_num=b.Series_num and a.Gridcode=b.Gridcode and a.habt=b.habt);
 quit;
```

CS1: Script to update plot habitat allocation table with Broad Habitats allocated from digitized 1978 area data

```
/*Script to update IBD78 table with Broad Habitats allocated from
digitized 1978 area data CMW 2.9.10*/
```

 # IMPORT SPREADSHEET CONTAINING UPDATES

```
PROC IMPORT OUT= WORK.BH78
            DATAFILE= "S:\PARR Section\LUS\projects\current
projects\CS1
978 Data Rescue\Final_data\Plot_BH_Alllocations.xls"
            DBMS=EXCEL REPLACE;
     SHEET="Final_BH_allocs78$";
     GETNAMES=YES;
     MIXED=NO;
     SCANTEXT=YES;
     USEDATE=YES;
     SCANTIME=YES;
RUN;
```

 # RUN UPDATES TO UPDATE EXISTING PLOT DATA WITH NEWLY DIGITISED HABITAT DATA

```
proc sql;
    update cs_veg.IBD78 a
    set BH78 = (select Broad_Habitat from work.BH78 b
           where a.Rep_id78=b.Rep_id)
    where a.Rep_id78 in (select Rep_id from work.BH78 b);
quit;

proc sql;
    update cs_veg.IBD78 a
    set MAPCODE78 = (select Mapcode_78 from work.BH78 b
           where a.Rep_id78=b.Rep_id)
    where a.Rep_id78 in (select Rep_id from work.BH78 b);
quit;
```

 # IMPORT SPREADSHEET CONTAINING UPDATES (WHERE HABITAT ALLOCATIONS ARE STILL MISSING, POST CHECKS)

```
PROC IMPORT OUT= WORK.BH78
            DATAFILE= "S:\PARR Section\LUS\projects\current
projects\CS1978 Data Rescue\Final_data\Plot_BH_Alllocations.xls"
            DBMS=EXCEL REPLACE;
```

```
      SHEET="Missing_plots_allocs_DB$";
      GETNAMES=YES;
      MIXED=NO;
      SCANTEXT=YES;
      USEDATE=YES;
      SCANTIME=YES;
RUN;
```

 # RUN UPDATES TO UPDATE EXISTING PLOT DATA WITH NEWLY DIGITISED HABITAT DATA

```
proc sql;
    update cs_veg.IBD78 a
    set BH78 = (select Broad_Habitat from work.BH78 b
           where a.Rep_id78=b.Rep_id)
    where a.Rep_id78 in (select Rep_id from work.BH78 b)and a.BH78
is null;
quit;

proc sql;
    update cs_veg.IBD78 a
    set MAPCODE78 = (select Mapcode_78 from work.BH78 b
           where a.Rep_id78=b.Rep_id)
    where a.Rep_id78 in (select Rep_id from work.BH78 b)and
a.MAPCODE78 is null;
quit;
```

Data Analyses Scripts

SPW4: Query to create table of the top 25 most frequently recorded ground flora species in pinewoods

 # QUERY GROUND FLORA TABLE BY SELECTING SPECIES AND COUNT (EXCLUDING BRYOPHYTES AND ORDERING BY DESCENDING COUNT

```
PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_SCOTS_PINE_1971_GROUND AS
   SELECT t1.BRC_NAME_FINAL,
          /* COUNT_of_BRC_NAME_FINAL */
            (COUNT(t1.BRC_NAME_FINAL)) AS COUNT_of_BRC_NAME_FINAL
      FROM BUNCE.SCOTS_PINE_1971_GROUND_FLORA t1
      WHERE t1.GROWTH_FORM NOT IN
           (
           'b'
           )
      GROUP BY t1.BRC_NAME_FINAL
      ORDER BY COUNT_of_BRC_NAME_FINAL DESC;
QUIT;
```

SPW5: Code to create charts showing tree diameter of Scots Pine (*Pinus sylvestris*) distribution in different woods

 # CLASSIFY TREE DIAMETERS INTO 5CM CLASSES AND SELECT *PINUS SYLVESTRIS*

```
PROC SQL;
```

```
   CREATE TABLE
WORK.QUERY_FOR_SCOTS_PINE_1971_T_0003(label="QUERY_FOR_SCOTS_PINE_1971_TREE_DATA")
AS
   SELECT t1.SITE_NO,
          t1.PLOT_NO,
          t1.TREE_NO,
          t1.TREE_TYPE,
          t1.SPECIES,
          t1.DBH,
          /* DBHclass */
            (CASE
             WHEN t1.DBH <=5
               THEN 1
            WHEN t1.DBH <=10
               THEN 2
            WHEN t1.DBH<=15
               THEN 3
            WHEN t1.DBH <=20
               THEN 4
            WHEN t1.DBH<=25
               THEN 5
            WHEN t1.DBH <=30
               THEN 6
            WHEN t1.DBH <=35
               THEN 7
            WHEN t1.DBH <=40
               THEN 8
            WHEN t1.DBH <=45
               THEN 9
            WHEN t1.DBH <=50
               THEN 10
            WHEN t1.DBH <=55
               THEN 11
            WHEN t1.DBH <=60
               THEN 12
            WHEN t1.DBH<=65
               THEN 13
            WHEN t1.DBH <=70
               THEN 14
            WHEN t1.DBH<=75
               THEN 15
            WHEN t1.DBH <=80
               THEN 16
            WHEN t1.DBH <=85
               THEN 17
            WHEN t1.DBH <=90
               THEN 18
            WHEN t1.DBH <=95
               THEN 19
            WHEN t1.DBH >95
               THEN 20
            ELSE 0
            END) AS DBHclass
      FROM BUNCE.SCOTS_PINE_1971_TREE_DATA t1
      WHERE t1.PLOT_NO <= 16 AND t1.SPECIES = 'Pinus sylvestris';
QUIT;
```

 # SUMMARISE NUMBER OF TREES IN EACH CLASS

```
PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_SCOTS_PINE_1971_T_0004(label="QUERY_FOR_SCOTS_PINE_1971_T") AS
   SELECT t1.SITE_NO,
          t1.PLOT_NO,
          t1.TREE_TYPE,
          t1.DBHclass,
          /* Count */
            (1) AS Count
```

```
      FROM WORK.QUERY_FOR_SCOTS_PINE_1971_T_0003 t1;
QUIT;
```

 # JOIN DATA TO TABLE CONTAINING SITE NAMES

```
PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_SCOTS_PINE_1971_T_0006(label="QUERY_FOR_SCOTS_PINE_1971_T") AS
   SELECT DISTINCT t1.SITE_NO,
          t2.SITE_NAME,
          t1.DBHclass,
          /* SUM_of_Count */
            (SUM(t1.Count)) AS SUM_of_Count
      FROM WORK.QUERY_FOR_SCOTS_PINE_1971_T_0004 t1
           INNER JOIN WORK.SCOTS_PINE_1971_SITES t2 ON (t1.SITE_NO = t2.SITEID)
      GROUP BY t1.SITE_NO,
               t2.SITE_NAME,
               t1.DBHclass;
QUIT;
```

 # CREATE BAR CHARTS PER SITE

```
/* -------------------------------------------------------------------
   Code generated by SAS Task

   Generated on: Thursday, July 23, 2020 at 12:20:54 PM
   By task: Bar Chart2

   Input Data: Local:WORK.QUERY_FOR_SCOTS_PINE_1971_T_0006
   Server:  Local
   ------------------------------------------------------------------- */

%_eg_conditional_dropds(WORK.SORTTempTableSorted);
/* -------------------------------------------------------------------
   Sort data set Local:WORK.QUERY_FOR_SCOTS_PINE_1971_T_0006
   ------------------------------------------------------------------- */

PROC SQL;
      CREATE VIEW WORK.SORTTempTableSorted AS
            SELECT T.DBHclass, T.SUM_of_Count, T.SITE_NO, T.SITE_NAME
      FROM WORK.QUERY_FOR_SCOTS_PINE_1971_T_0006 as T
;
QUIT;
%_sas_pushchartsize(480,150);
      PATTERN1 COLOR=BLACK;
      PATTERN2 COLOR = _STYLE_;
      PATTERN3 COLOR = _STYLE_;
      PATTERN4 COLOR = _STYLE_;
      PATTERN5 COLOR = _STYLE_;
      PATTERN6 COLOR = _STYLE_;
      PATTERN7 COLOR = _STYLE_;
      PATTERN8 COLOR = _STYLE_;
      PATTERN9 COLOR = _STYLE_;
      PATTERN10 COLOR = _STYLE_;
      PATTERN11 COLOR = _STYLE_;
      PATTERN12 COLOR = _STYLE_;
Axis1
      STYLE=1
      WIDTH=1
      ORDER=(0 TO 100 BY 10)
      MINOR=NONE
      LABEL=( FONT='Arial'  ANGLE=90 "Number of trees")

;
Axis2
      STYLE=1
      WIDTH=1
      LABEL=( FONT='Arial'   "5cm diameter class")
```

```
;
TITLE;
FOOTNOTE;
PROC GCHART DATA=WORK.SORTTempTableSorted
;
       VBAR
        DBHclass
 /
       SUMVAR=SUM_of_Count
       CLIPREF
              SPACE=0
NOFRAME       MIDPOINTS=1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
       TYPE=SUM
       COUTLINE=SAME
       RAXIS=AXIS1
       MAXIS=AXIS2
;
       BY SITE_NO SITE_NAME;
```

CS2: Code to create estimates of Broad Habitat area features in upland areas

 # IMPORT BROAD HABITAT DATA FOR CS 2007

```
/*   START OF NODE: Import Data (CS007_BroadHabitats_Stock.csv)   */

DATA WORK.CS007_BroadHabitats_Stock;
    LENGTH
        YEAR                8
        LAND_CLASS        $ 4
        BROAD_HABITAT       8
        BROAD_HABITAT_NAME $ 35
        LAND_CLASS_AREA     8
        MEAN_ESTIMATE       8
        LOWER_ESTIMATE      8
        UPPER_ESTIMATE      8 ;
    FORMAT
        YEAR             BEST4.
        LAND_CLASS       $CHAR4.
        BROAD_HABITAT    BEST2.
        BROAD_HABITAT_NAME $CHAR35.
        LAND_CLASS_AREA  BEST7.
        MEAN_ESTIMATE    BEST12.
        LOWER_ESTIMATE   BEST12.
        UPPER_ESTIMATE   BEST11. ;
    INFORMAT
        YEAR             BEST4.
        LAND_CLASS       $CHAR4.
        BROAD_HABITAT    BEST2.
        BROAD_HABITAT_NAME $CHAR35.
        LAND_CLASS_AREA  BEST7.
        MEAN_ESTIMATE    BEST12.
        LOWER_ESTIMATE   BEST12.
        UPPER_ESTIMATE   BEST11. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG12076\CS007_BroadHabitats_Stock-
c74f54b8791c485c9b05b26f8dc99e45.txt'
        LRECL=90
        ENCODING="WLATIN1"
        TERMSTR=CRLF
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        YEAR             : ?? BEST4.
        LAND_CLASS       : $CHAR4.
```

```
        BROAD_HABITAT    : ?? BEST2.
        BROAD_HABITAT_NAME : $CHAR35.
        LAND_CLASS_AREA  : ?? COMMA7.
        MEAN_ESTIMATE    : ?? COMMA12.
        LOWER_ESTIMATE   : ?? COMMA12.
        UPPER_ESTIMATE   : ?? COMMA11. ;
RUN;
```

 # FILTER HABITAT DATA FROM COUNTRYSIDE SURVEY SITES THAT ARE LOCATED IN UPLANDS

```
/*   START OF NODE: Query Builder2   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_SQUARES_FILE_ALL__0001(label="QUERY_FOR_SQUARES_FILE_ALL_LC") AS
   SELECT t1.ENV_ZONE_2007,
          t1.LC07,
          t1.LC07_NUM,
          t2.YEAR,
          t2.LAND_CLASS,
          t2.BROAD_HABITAT,
          t2.BROAD_HABITAT_NAME,
          t2.LAND_CLASS_AREA,
          t2.MEAN_ESTIMATE,
          t2.LOWER_ESTIMATE,
          t2.UPPER_ESTIMATE,
          t1.COUNTRY
      FROM WORK.QUERY_FOR_SQUARES_FILE_ALL_LC t1
           LEFT JOIN WORK.CS007_BROADHABITATS_STOCK t2 ON (t1.LC07 =
t2.LAND_CLASS);
QUIT;
```

 # SUM ESTIMATES BY COUNTRY: OUTPUT

```
/*   START OF NODE: Query Builder20   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_SQUARES_FILE_ALL_0005(label="QUERY_FOR_SQUARES_FILE_ALL") AS
   SELECT t1.COUNTRY,
          t1.BROAD_HABITAT,
          t1.BROAD_HABITAT_NAME,
          /* SUM_of_LOWER_ESTIMATE */
            (SUM(t1.LOWER_ESTIMATE)) FORMAT=BEST12. AS SUM_of_LOWER_ESTIMATE,
          /* SUM_of_MEAN_ESTIMATE */
            (SUM(t1.MEAN_ESTIMATE)) FORMAT=BEST12. AS SUM_of_MEAN_ESTIMATE,
          /* SUM_of_UPPER_ESTIMATE */
            (SUM(t1.UPPER_ESTIMATE)) FORMAT=BEST11. AS SUM_of_UPPER_ESTIMATE
      FROM WORK.QUERY_FOR_SQUARES_FILE_ALL__0001 t1
      GROUP BY t1.COUNTRY,
               t1.BROAD_HABITAT,
               t1.BROAD_HABITAT_NAME;
QUIT;
```

CS3: Code to create estimates of linear features in upland areas

 # FILTER LAND CLASSES FROM COUNTRYSIDE SURVEY THAT ARE LOCATED IN UPLANDS

```
/*   START OF NODE: Query Builder   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_SQUARES_FILE_ALL_LC AS
   SELECT DISTINCT t1.ENV_ZONE_2007,
          t1.LC07,
          t1.LC07_NUM,
```

```
            t1.COUNTRY
      FROM CS_SQS.SQUARES_FILE_ALL_LC t1
      WHERE t1.ENV_ZONE_2007 IN
           (
           3,
           5,
           6,
           9
           )
      ORDER BY t1.LC07_NUM,
               t1.ENV_ZONE_2007;
QUIT;
```

 # FILTER SITES FROM COUNTRYSIDE SURVEY NATIONAL ESTIMATES FOR LINEAR FEATURES FOR UPLANDS

```
/*   START OF NODE: Query Builder12   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_SQUARES_FILE_ALL__0002(label="QUERY_FOR_SQUARES_FILE_ALL_LC") AS
   SELECT t1.ENV_ZONE_2007,
          t1.LC07,
          t1.LC07_NUM,
          t1.COUNTRY,
          t2._Estimate_type,
          t2._land_class,
          t2._attribute,
          t2._survey,
          t2._n,
          t2._mean,
          t2._se,
          t2._pctl0_5,
          t2._pctl2_5,
          t2._pctl50,
          t2._pctl97_5,
          t2._pctl99_5,
          t2._area
      FROM WORK.QUERY_FOR_SQUARES_FILE_ALL_LC t1
           LEFT JOIN CS_SEC.LC_ESTIMATES_LN_1984199019982007 t2 ON (t1.LC07 =
t2._land_class);
QUIT;
```

 # FILTER SITES FROM COUNTRYSIDE SURVEY NATIONAL ESTIMATES FOR LINEAR FEATURES SURVEYED IN 2007

```
/*   START OF NODE: Query Builder13   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_SQUARES_FILE_ALL AS
   SELECT t1.ENV_ZONE_2007,
          t1.LC07,
          t1.LC07_NUM,
          t1.COUNTRY,
          t1._Estimate_type,
          t1._land_class,
          t1._attribute,
          t1._survey,
          t1._n,
          t1._mean,
          t1._se,
          t1._pctl0_5,
          t1._pctl2_5,
          t1._pctl50,
          t1._pctl97_5,
          t1._pctl99_5,
          t1._area
      FROM WORK.QUERY_FOR_SQUARES_FILE_ALL__0002 t1
      WHERE t1._survey = '        2007';
QUIT;
```

 # FILTER SITES FROM COUNTRYSIDE SURVEY NATIONAL ESTIMATES FOR LINEAR FEATURES ABOVE (PARAMETRIC MEANS)

```
/*   START OF NODE: Query Builder14   */

PROC SQL;
   CREATE TABLE WORK.LINES_PARA_NEW(label="QUERY_FOR_SQUARES_FILE_ALL") AS
   SELECT t1.ENV_ZONE_2007,
          t1.LC07,
          t1.LC07_NUM,
          t1.COUNTRY,
          t1._Estimate_type,
          t1._land_class,
          t1._attribute,
          t1._survey,
          t1._area,
          /* MEAN */
            ((t1._mean * t1._area)/1000000) AS MEAN
      FROM WORK.QUERY_FOR_SQUARES_FILE_ALL t1
      WHERE t1._Estimate_type = 'Parametric New';
QUIT;
```

 # SUMMARISE DATA FROM COUNTRYSIDE SURVEY NATIONAL ESTIMATES FOR LINEAR FEATURES BY COUNTRY

```
/*   START OF NODE: Query Builder16   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_LINES_PARA_NEW AS
   SELECT t1.COUNTRY,
          t1._attribute,
          /* SUM_of_MEAN */
            (SUM(t1.MEAN)) AS SUM_of_MEAN
      FROM WORK.LINES_PARA_NEW t1
      GROUP BY t1.COUNTRY,
               t1._attribute
      ORDER BY t1.COUNTRY,
               t1._attribute;
QUIT;
```

 # FILTER SITES FROM COUNTRYSIDE SURVEY NATIONAL ESTIMATES FOR LINEAR FEATURES ABOVE (CONSISTENT MODEL MEANS)

```
/*   START OF NODE: Query Builder15   */


PROC SQL;
   CREATE TABLE WORK.LINES_CONS(label="QUERY_FOR_SQUARES_FILE_ALL") AS
   SELECT t1.ENV_ZONE_2007,
          t1.LC07,
          t1.LC07_NUM,
          t1.COUNTRY,
          t1._Estimate_type,
          t1._land_class,
          t1._attribute,
          t1._survey,
          t1._n,
          t1._mean,
          t1._se,
          t1._pctl0_5,
          t1._pctl2_5,
          t1._pctl50,
          t1._pctl97_5,
          t1._pctl99_5,
          t1._area
      FROM WORK.QUERY_FOR_SQUARES_FILE_ALL t1
      WHERE t1._Estimate_type = 'Consistent';
QUIT;
```

CALCULATE UPPER AND LOWER ESTIMATES FOR COUNTRYSIDE SURVEY LINEAR FEATURES ABOVE

```
/*   START OF NODE: Query Builder17   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_LINES_CONS AS
   SELECT t1.ENV_ZONE_2007,
          t1.LC07,
          t1.LC07_NUM,
          t1.COUNTRY,
          t1._Estimate_type,
          t1._land_class,
          t1._attribute,
          t1._survey,
          t1._n,
          t1._mean,
          t1._se,
          t1._pctl0_5,
          t1._pctl2_5,
          t1._pctl50,
          t1._pctl97_5,
          t1._pctl99_5,
          t1._area,
          /* UPPER */
            (t1._pctl97_5 * t1._area /1000000) AS UPPER,
          /* LOWER */
            (t1._pctl2_5 * t1._area /1000000) AS LOWER
      FROM WORK.LINES_CONS t1;
QUIT;
```

CREATE TABLE OF MEAN ESTIMATES, AND UPPER AND LOWER ESTIMATES FOR COUNTRYSIDE SURVEY LINEAR FEATURES ABOVE

```
/*   START OF NODE: Query Builder18   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_LINES_CONS_0000(label="QUERY_FOR_LINES_CONS") AS
   SELECT t1.COUNTRY,
          t1._attribute,
          /* SUM_of_UPPER */
            (SUM(t1.UPPER)) AS SUM_of_UPPER,
          /* SUM_of_LOWER */
            (SUM(t1.LOWER)) AS SUM_of_LOWER
      FROM WORK.QUERY_FOR_LINES_CONS t1
      GROUP BY t1.COUNTRY,
               t1._attribute;
QUIT;
```

CREATE TABLE OF MEAN ESTIMATES, AND UPPER AND LOWER ESTIMATES FOR COUNTRYSIDE SURVEY LINEAR FEATURES ABOVE

```
/*   START OF NODE: Query Builder19   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_LINES_PARA_NEW_0000(label="QUERY_FOR_LINES_PARA_NEW") AS
   SELECT t1.COUNTRY,
          t1._attribute,
          t2.SUM_of_LOWER,
          t1.SUM_of_MEAN,
          t2.SUM_of_UPPER
      FROM WORK.QUERY_FOR_LINES_PARA_NEW t1, WORK.QUERY_FOR_LINES_CONS_0000 t2
      WHERE (t1.COUNTRY = t2.COUNTRY AND t1._attribute = t2._attribute);
QUIT;
```

 # CREATE TABLE OF MEAN ESTIMATES, AND UPPER AND LOWER ESTIMATES FOR COUNTRYSIDE SURVEY LINEAR FEATURES ABOVE, GROUP BY BOUNDARY TYPE

```
/*   START OF NODE: Query Builder47   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_LINES_PARA_NEW_0001(label="QUERY_FOR_LINES_PARA_NEW") AS
   SELECT t1._attribute,
          /* SUM_of_SUM_of_MEAN */
            (SUM(t1.SUM_of_MEAN)) AS SUM_of_SUM_of_MEAN
      FROM WORK.QUERY_FOR_LINES_PARA_NEW_0000 t1
      GROUP BY t1._attribute;
QUIT;
```

CS4: Code to summarise estimates of linear features by type and by country

 # CREATE SUMMARY OF LAND CLASSES FROM COUNTRYSIDE SURVEY THAT ARE LOCATED IN UPLANDS, BY COUNTRY

```
/*   START OF NODE: Query Builder23   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_SQUARES_FILE_ALL__0003(label="QUERY_FOR_SQUARES_FILE_ALL_LC") AS
   SELECT DISTINCT t1.COUNTRY,
          t1.LC07,
          t1.LC07_NUM,
          t1.ENV_ZONE_2007
      FROM CS_SQS.SQUARES_FILE_ALL_LC t1;
QUIT;
```

 # SELECT MEAN ESTIMATES DATA FOR BOUNDARY FEATURES IN UPLAND SITES

```
/*   START OF NODE: Query Builder21   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_LC_ESTIMATES_LN_198419(label="QUERY_FOR_LC_ESTIMATES_LN_198419901998
") AS
   SELECT t1._Estimate_type,
          t1._attribute,
          t1._survey,
          t1._mean,
          t1._area,
          t2.COUNTRY,
          t2.LC07,
          t2.LC07_NUM,
          t2.ENV_ZONE_2007
      FROM CS_SEC.LC_ESTIMATES_LN_1984199019982007 t1
           LEFT JOIN WORK.QUERY_FOR_SQUARES_FILE_ALL__0003 t2 ON (t1._land_class =
t2.LC07)
      WHERE t1._survey = '        2007' AND t1._Estimate_type = 'Parametric New';
QUIT;
```

 # CALCULATE MEAN ESTIMATES FOR LINEAR FEATURES IN KM

```
/*   START OF NODE: Query Builder22   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_LC_ESTIMATES_LN_19 AS
   SELECT t1.COUNTRY,
          t1._attribute,
          t1._survey,
          t1._mean,
          t1._area,
          /* MEAN */
```

```
            (t1._mean * t1._area / 1000000) AS MEAN
      FROM WORK.QUERY_FOR_LC_ESTIMATES_LN_198419 t1;
QUIT;
```

 # OUTPUT: MEAN ESTIMATES FOR LINEAR FEATURES IN KM, BY TYPE AND COUNTRY

```
/*   START OF NODE: Query Builder24   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_LC_ESTIMATES_LN_1_0000(label="QUERY_FOR_LC_ESTIMATES_LN_19") AS
   SELECT t1.COUNTRY,
          t1._attribute,
          /* SUM_of_MEAN */
            (SUM(t1.MEAN)) AS SUM_of_MEAN
      FROM WORK.QUERY_FOR_LC_ESTIMATES_LN_19 t1
      GROUP BY t1.COUNTRY,
               t1._attribute;
QUIT;
```

CS5: Code to summarise plots in upland areas

 # CREATE TABLE OF ENVIRONMENTAL ZONES

```
/*   START OF NODE: Query Builder1   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_SQUARES_FILE_ALL__0000(label="QUERY_FOR_SQUARES_FILE_ALL_LC") AS
   SELECT DISTINCT t1.ENV_ZONE_2007,
          t1.ENV_ZONE_1990
      FROM CS_SQS.SQUARES_FILE_ALL_LC t1
      ORDER BY t1.ENV_ZONE_2007;
QUIT;
```

 # SELECT VEGETATION PLOT CHARACTERISTICS FROM UPLAND SITES

```
/*   START OF NODE: Query Builder4   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_IBD07 AS
   SELECT t1.REP_ID07,
          t1.SERIES_NUM,
          t1.AMALG_PTYPE,
          t1.PLOT_TYPE,
          t1.CVS07,
          t1.AVC07,
          t1.BH07,
          t1.PH07,
          t2.LC07,
          t2.LC07_NUM
      FROM CS_VEG.IBD07 t1
           LEFT JOIN CS_SQS.SQUARES_FILE_ALL_LC t2 ON (t1.SERIES_NUM =
t2.SERIES_NUM);
QUIT;
```

 # SELECT VEGETATION PLOT CHARACTERISTICS FROM UPLAND SITES

```
/*   START OF NODE: Query Builder5   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_IBD07_0000(label="QUERY_FOR_IBD07") AS
   SELECT t1.REP_ID07,
          t1.SERIES_NUM,
```

```
         t1.AMALG_PTYPE,
         t1.PLOT_TYPE,
         t1.CVS07,
         t1.AVC07,
         t1.BH07,
         t1.PH07,
         t1.LC07,
         t1.LC07_NUM,
         t2.COUNTRY
     FROM WORK.QUERY_FOR_IBD07 t1
          RIGHT JOIN WORK.QUERY_FOR_SQUARES_FILE_ALL_LC t2 ON (t1.LC07 = t2.LC07);
QUIT;
```

 # SUMMARISE COUNTRYSIDE VEGETATION SYSTEM GROUPS FROM VEGETATION PLOTS FROM UPLAND SITES: OUTPUT

```
/*   START OF NODE: Query Builder6   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_IBD07_0001(label="QUERY_FOR_IBD07") AS
   SELECT t1.COUNTRY,
          t1.CVS07,
          /* COUNT_of_CVS07 */
            (COUNT(t1.CVS07)) AS COUNT_of_CVS07
      FROM WORK.QUERY_FOR_IBD07_0000 t1
      GROUP BY t1.COUNTRY,
               t1.CVS07
      ORDER BY t1.COUNTRY,
               COUNT_of_CVS07,
               t1.CVS07;
QUIT;
```

 # SUMMARISE AGGREGATED COUNTRYSIDE VEGETATION SYSTEM GROUPS FROM VEGETATION PLOTS FROM UPLAND SITES: OUTPUT

```
/*   START OF NODE: Query Builder7   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_IBD07_0002(label="QUERY_FOR_IBD07") AS
   SELECT t1.COUNTRY,
          t1.AVC07,
          t2.AGG_CLASS_DESCRIPTION,
          /* COUNT_of_AVC07 */
            (COUNT(t1.AVC07)) AS COUNT_of_AVC07
      FROM WORK.QUERY_FOR_IBD07_0000 t1
           LEFT JOIN CS_SQS.CVS_AGG_CLASS_DESC t2 ON (t1.AVC07 = t2.AGG_CLASS_NUM)
      GROUP BY t1.COUNTRY,
               t1.AVC07,
               t2.AGG_CLASS_DESCRIPTION;
QUIT;
```

CS6: Code to Summarise vegetation plots

 # SELECT VEGETATION PLOT DATA FROM 2007 WHERE PLOT WAS RECORDED

```
/*   START OF NODE: Query Builder8   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_VEGETATION_PLOTS AS
   SELECT t1.VEG_PLOTS_ID,
          t1.SQUARE,
          t1.PLOT_TYPE,
          t1.PLOT_NUMBER,
          t1.REP_ID,
          t1.RECORDED,
          t2.BRC_NUMBER,
```

```
          t2.TOTAL_COVER
      FROM CS_VEG.VEGETATION_PLOTS_2007 t1
           LEFT JOIN CS_VEG.VEGETATION_PLOT_SPECIES_2007 t2 ON (t1.VEG_PLOTS_ID =
t2.VEG_PLOTS_ID)
      WHERE t1.RECORDED IN
           (
           1,
           3,
           4
           );
QUIT;
```

 # SELECT VEGETATION PLOT SPECIES DATA FROM 2007

```
/*   START OF NODE: Query Builder9   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_VEGETATION_PLOTS_0000(label="QUERY_FOR_VEGETATION_PLOTS") AS
   SELECT t1.VEG_PLOTS_ID,
          t1.SQUARE,
          t1.PLOT_TYPE,
          t1.PLOT_NUMBER,
          t1.REP_ID,
          t1.RECORDED,
          t1.BRC_NUMBER,
          t2.LC07,
          t2.LC07_NUM,
          t1.TOTAL_COVER
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS t1
           LEFT JOIN WORK.QUERY_FOR_IBD07 t2 ON (t1.REP_ID = t2.REP_ID07);
QUIT;
```

 # FILTER VEGETATION PLOT SPECIES DATA FROM 2007 FOR THE UPLANDS

```
/*   START OF NODE: Query Builder10   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_VEGETATION_PLOTS_0001(label="QUERY_FOR_VEGETATION_PLOTS") AS
   SELECT t1.SQUARE,
          t1.PLOT_TYPE,
          t1.PLOT_NUMBER,
          t1.REP_ID,
          t1.RECORDED,
          t1.BRC_NUMBER,
          t1.LC07,
          t1.LC07_NUM,
          t2.BRC_NAMES,
          t3.COUNTRY,
          t1.TOTAL_COVER
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS_0000 t1
           LEFT JOIN CS_VEG.LUS_SP_LIB_AND_TRAITS t2 ON (t1.BRC_NUMBER =
t2.BRC_NUMBER)
           RIGHT JOIN WORK.QUERY_FOR_SQUARES_FILE_ALL_LC t3 ON (t1.LC07 = t3.LC07);
QUIT;
```

 # CALCULATE MEAN, TOTALS AND COUNT OF SPECIES COVERS FROM 2007 FOR THE UPLANDS

```
/*   START OF NODE: Query Builder11   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_VEGETATION_PLOTS_0002(label="QUERY_FOR_VEGETATION_PLOTS") AS
   SELECT DISTINCT t1.COUNTRY,
          t1.BRC_NAMES,
          /* COUNT_of_BRC_NAMES */
```

```
            (COUNT(t1.BRC_NAMES)) AS COUNT_of_BRC_NAMES,
          /* SUM_of_TOTAL_COVER */
            (SUM(t1.TOTAL_COVER)) FORMAT=6. AS SUM_of_TOTAL_COVER,
          /* MEAN_of_TOTAL_COVER */
            (MEAN(t1.TOTAL_COVER)) FORMAT=6. AS MEAN_of_TOTAL_COVER
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS_0001 t1
      GROUP BY t1.COUNTRY,
               t1.BRC_NAMES
      ORDER BY t1.COUNTRY,
               COUNT_of_BRC_NAMES DESC;
QUIT;
```

 # FILTER ABOVE FOR ENGLAND ONLY: OUTPUT

```
/*   START OF NODE: Query Builder27   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_VEGETATION_PLOTS_0007(label="QUERY_FOR_VEGETATION_PLOTS") AS
   SELECT t1.COUNTRY,
          t1.BRC_NAMES,
          t1.COUNT_of_BRC_NAMES,
          t1.SUM_of_TOTAL_COVER,
          t1.MEAN_of_TOTAL_COVER
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS_0002 t1
      WHERE t1.COUNTRY = 'ENG';
QUIT;
```

 # FILTER ABOVE FOR WALES ONLY: OUTPUT

```
/*   START OF NODE: Query Builder28   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_VEGETATION_PLOTS_0008(label="QUERY_FOR_VEGETATION_PLOTS") AS
   SELECT t1.COUNTRY,
          t1.BRC_NAMES,
          t1.COUNT_of_BRC_NAMES,
          t1.SUM_of_TOTAL_COVER
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS_0002 t1
      WHERE t1.COUNTRY = 'WAL';
QUIT;
```

 # FILTER ABOVE FOR SCOTLAND ONLY: OUTPUT

```
/*   START OF NODE: Query Builder29   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_VEGETATION_PLOTS_0009(label="QUERY_FOR_VEGETATION_PLOTS") AS
   SELECT t1.COUNTRY,
          t1.BRC_NAMES,
          t1.COUNT_of_BRC_NAMES,
          t1.SUM_of_TOTAL_COVER
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS_0002 t1
      WHERE t1.COUNTRY = 'SCO';
QUIT;
```

CS7: Summarise Countryside Survey vegetation plots by country

 # SELECT COUNTRYSIDE SURVEY 2007 VEGETATION PLOTS LOCATED IN UPLAND ZONES BY JOINING TO LAND CLASS TABLE

```
/*   START OF NODE: Query Builder43   */
```

```
PROC SQL;
   CREATE TABLE 
WORK.QUERY_FOR_VEGETATION_PLOTS_0017(label="QUERY_FOR_VEGETATION_PLOTS") AS 
   SELECT DISTINCT t1.SQUARE, 
          t1.PLOT_TYPE, 
          t1.PLOT_NUMBER, 
          t1.REP_ID, 
          t1.RECORDED, 
          t2.ENV_ZONE_2007, 
          t2.COUNTRY, 
          t2.LC07_NUM, 
          t2.LC07
      FROM CS_VEG.VEGETATION_PLOTS_2007 t1
           LEFT JOIN CS_SQS.SQUARES_FILE_ALL_LC t2 ON (t1.SQUARE = t2.SERIES_NUM)
      WHERE t1.RECORDED IN 
           (
           1,
           3,
           4
           ) AND t2.ENV_ZONE_2007 IN 
           (
           3,
           5,
           6,
           9
           );
QUIT;
```

 # COUNT COUNTRYSIDE SURVEY 2007 VEGETATION PLOTS LOCATED IN UPLAND ZONES

```
/*   START OF NODE: Query Builder44   */

PROC SQL;
   CREATE TABLE 
WORK.QUERY_FOR_VEGETATION_PLOTS_0018(label="QUERY_FOR_VEGETATION_PLOTS") AS 
   SELECT t1.ENV_ZONE_2007, 
          t1.PLOT_TYPE, 
          /* COUNT_of_PLOT_TYPE */
            (COUNT(t1.PLOT_TYPE)) AS COUNT_of_PLOT_TYPE
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS_0017 t1
      WHERE t1.ENV_ZONE_2007 IN 
           (
           3,
           5,
           6,
           9
           )
      GROUP BY t1.ENV_ZONE_2007,
               t1.PLOT_TYPE;
QUIT;
```

 # COUNT COUNTRYSIDE SURVEY 2007 VEGETATION PLOTS LOCATED IN UPLAND ZONES, SUMMARISE BY COUNTRY

```
/*   START OF NODE: Query Builder45   */

PROC SQL;
   CREATE TABLE 
WORK.QUERY_FOR_VEGETATION_PLOTS_0019(label="QUERY_FOR_VEGETATION_PLOTS") AS 
   SELECT t1.COUNTRY, 
          /* COUNT_of_PLOT_TYPE */
            (COUNT(t1.PLOT_TYPE)) AS COUNT_of_PLOT_TYPE
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS_0017 t1
```

```
      GROUP BY t1.COUNTRY;
QUIT;
```

 # COUNT COUNTRYSIDE SURVEY 2007 VEGETATION PLOTS (X plots only) LOCATED IN UPLAND ZONES, SUMMARISE BY COUNTRY

```
/*   START OF NODE: Query Builder46   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_VEGETATION_PLOTS_001A(label="QUERY_FOR_VEGETATION_PLOTS") AS
   SELECT t1.COUNTRY,
          t1.PLOT_TYPE,
          /* COUNT_of_PLOT_TYPE */
            (COUNT(t1.PLOT_TYPE)) AS COUNT_of_PLOT_TYPE
      FROM WORK.QUERY_FOR_VEGETATION_PLOTS_0017 t1
      WHERE t1.PLOT_TYPE = 'X'
      GROUP BY t1.COUNTRY,
               t1.PLOT_TYPE;
QUIT;
```

CS8: Code to create estimates of grazed areas in the uplands

 # JOIN COUNTRYSIDE SURVEY HABITAT DATA TO FILE CONTAINING LIST OF UPLAND SITES

```
/*   START OF NODE: Query Builder1   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_AREA AS
   SELECT t1.SERIES_NUM,
          t1.SCPTDATA_ID,
          t1.BROAD_HABITAT,
          t1.BH98_ACCURACY,
          t1.VISIT_STATUS,
          t1.REASON_FOR_CHANGE,
          t1.POLYGON_AREA,
          t1.LENGTH,
          t2.SCPTDATA_ID AS SCPTDATA_ID1,
          t2.SERIES_NUM AS SERIES_NUM1,
          t2.LUSE,
          t2.HABT_CODE,
          t2.SPECIES,
          t2.SPECIES_COVER,
          t2.PRIMARY_QUALIFIER,
          t2.STRUCTURE_USE,
          t2.PHYSIOGRAPHY_COVER,
          t2.ROAD_VERGE_A,
          t2.ROAD_VERGE_B,
          t2.MODAL_DBH,
          t2.AREAP
      FROM CS07_ANA.AREA t1
           LEFT JOIN CS07_ANA.AREA_ATTRIBUTE t2 ON (t1.SCPTDATA_ID =
t2.SCPTDATA_ID);
QUIT;
```

 # SELECT AREAS FROM COUNTRYSIDE SURVEY HABITAT DATA CONTAINING GRAZING ANIMALS (Cattle, Deer, Grouse, Sheep)

```
/*   START OF NODE: Query Builder   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_HABITATS AS
```

```
   SELECT t1.OBJECTID,
          t1.HABITAT_NAME,
          t1.HABT_CODE
      FROM CS07_ANA.HABITATS t1
      WHERE t1.HABITAT_NAME IN
           (
           'Cattle',
           'Deer',
           'Grouse',
           'Sheep'
           );
QUIT;
```

 # FILTER AREAS FROM COUNTRYSIDE SURVEY HABITAT DATA THAT WERE SURVEYED IN 2007

```
/*   START OF NODE: Query Builder2   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_AREA_0000(label="QUERY_FOR_AREA") AS
   SELECT t1.SERIES_NUM,
          t1.SCPTDATA_ID,
          t1.BROAD_HABITAT,
          t1.VISIT_STATUS,
          t1.POLYGON_AREA,
          t2.HABITAT_NAME
      FROM WORK.QUERY_FOR_AREA t1
           LEFT JOIN WORK.QUERY_FOR_HABITATS t2 ON (t1.HABT_CODE = t2.HABT_CODE)
      WHERE t1.VISIT_STATUS NOT = 3 AND t2.HABITAT_NAME NOT IS MISSING
      ORDER BY t1.SERIES_NUM;
QUIT;
```

 # FILTER AREAS FROM COUNTRYSIDE SURVEY HABITAT DATA LOCATED IN UPLANDS

```
/*   START OF NODE: Query Builder3   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_AREA_0001(label="QUERY_FOR_AREA") AS
   SELECT DISTINCT t1.SERIES_NUM,
          t1.SCPTDATA_ID,
          t1.BROAD_HABITAT,
          t1.VISIT_STATUS,
          t1.POLYGON_AREA,
          t2.ENV_ZONE_2007
      FROM WORK.QUERY_FOR_AREA_0000 t1
           LEFT JOIN CS_SQS.SQUARES_FILE_ALL_LC t2 ON (t1.SERIES_NUM =
t2.SERIES_NUM)
      WHERE t2.ENV_ZONE_2007 IN
           (
           3,
           5,
           6,
           9
           );
QUIT;

/*   START OF NODE: Query Builder4   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_AREA_0002(label="QUERY_FOR_AREA") AS
   SELECT DISTINCT t1.ENV_ZONE_2007,
          /* SUM_of_POLYGON_AREA */
            (SUM(t1.POLYGON_AREA)) AS SUM_of_POLYGON_AREA
      FROM WORK.QUERY_FOR_AREA_0001 t1
      GROUP BY t1.ENV_ZONE_2007;
QUIT;
```

 # FILTER SITES FROM COUNTRYSIDE SURVEY HABITAT DATA THAT WERE SURVEYED IN 2007

```
/*   START OF NODE: Query Builder5   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_SQUARES_FILE_ALL_LC AS
   SELECT t1.SERIES_NUM,
          t1.ENV_ZONE_2007,
          t2.CS2007
      FROM CS_SQS.SQUARES_FILE_ALL_LC t1
           LEFT JOIN CS_SQS.SURVEY_SQUARE_HISTORY t2 ON (t1.SERIES_NUM =
t2.SERIES_NUM)
      WHERE t2.CS2007 = 1;
QUIT;
```

 # COUNT NUMBER OF SITES FROM COUNTRYSIDE SURVEY HABITAT DATA THAT WERE SURVEYED IN 2007

```
/*   START OF NODE: Query Builder6   */

PROC SQL;
   CREATE TABLE
WORK.QUERY_FOR_SQUARES_FILE_ALL__0000(label="QUERY_FOR_SQUARES_FILE_ALL_LC") AS
   SELECT t1.ENV_ZONE_2007,
          /* COUNT_of_SERIES_NUM */
            (COUNT(t1.SERIES_NUM)) AS COUNT_of_SERIES_NUM
      FROM WORK.QUERY_FOR_SQUARES_FILE_ALL_LC t1
      GROUP BY t1.ENV_ZONE_2007;
QUIT;
```

 # CALCULATE GRAZED AREAS AS A PERCENTAGE OF SURVYED AREAS

```
/*   START OF NODE: Query Builder7   */

PROC SQL;
   CREATE TABLE WORK.QUERY_FOR_AREA_0004(label="QUERY_FOR_AREA") AS
   SELECT t1.ENV_ZONE_2007,
          t1.SUM_of_POLYGON_AREA,
          t2.COUNT_of_SERIES_NUM AS No_sqs_n_zone,
          /* Area_persq */
            (t1.SUM_of_POLYGON_AREA / t2.COUNT_of_SERIES_NUM) AS Area_persq,
          /* Prop */
            (((t1.SUM_of_POLYGON_AREA / t2.COUNT_of_SERIES_NUM) / 1000000)*100) AS
Prop
      FROM WORK.QUERY_FOR_AREA_0002 t1
           LEFT JOIN WORK.QUERY_FOR_SQUARES_FILE_ALL__0000 t2 ON (t1.ENV_ZONE_2007
= t2.ENV_ZONE_2007);
QUIT;
```

 # IMPORT ENVRONNMENTAL ZONE DATA

```
/*   START OF NODE: Import Data (Zone_areas.xlsx[Sheet1])   */

DATA WORK.Zone_areas;
    LENGTH
        EZ                 8
        Zone             $ 44
        Area               8 ;
    FORMAT
        EZ               BEST12.
        Zone             $CHAR44.
        Area             BEST12. ;
    INFORMAT
        EZ               BEST12.
        Zone             $CHAR44.
        Area             BEST12. ;
    INFILE 'C:\Users\clamw\AppData\Local\Temp\SEG13692\Zone_areas-
3c0478b80e8f4318ac327332605a4626.txt'
        LRECL=63
        ENCODING="WLATIN1"
        TERMSTR=CRLF
```

```
        DLM='7F'x
        MISSOVER
        DSD ;
    INPUT
        EZ               : BEST32.
        Zone             : $CHAR44.
        Area             : BEST32. ;
RUN;
```

 # OUTPUT: AREAS OF GRAZED LAND in ‘000ha

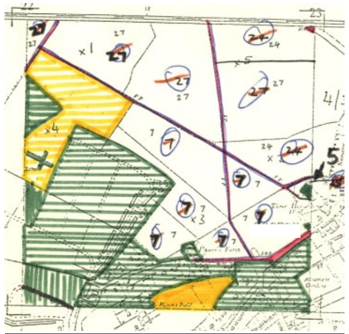
```
/*   START OF NODE: Query Builder8   */

PROC SQL;
   CREATE TABLE WORK.Area_by_animalzone(label="QUERY_FOR_AREA") AS
   SELECT t1.ENV_ZONE_2007,
          t1.SUM_of_POLYGON_AREA,
          t1.No_sqs_n_zone,
          t1.Area_persq,
          t1.Prop,
          t2.Area AS ZOneArea,
          /* Calculation */
            (t1.Prop * t2.Area) AS Calculation,
          /* Calculation1 */
            ((t1.Prop * t2.Area) / 1000000) AS Calculation1,
          /* Area_000ha */
            (((t1.Prop * t2.Area) / 1000000) / 1000) AS Area_000ha
      FROM WORK.QUERY_FOR_AREA_0004 t1
           LEFT JOIN WORK.ZONE_AREAS t2 ON (t1.ENV_ZONE_2007 = t2.EZ);
QUIT;
```

CS9: CS1978 digitizing protocol

Countryside Survey 1978 – Habitat Maps
Digitization Project

1. Background to data

Habitat and landscape features have been mapped and described, using field codes, for 256 individual 1km x 1km sample squares. All the maps are still in paper form and require digitizing into a single feature class.



Example of a mapped square

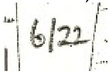
2. Digitizing Guidelines

2.1 Data Structure

- Areas
    - All polygons (representing areas) from each sample square are to be digitized into one feature class named Areas_1978. This will contain all data for all squares.
    - Each polygon should have a unique ID.
    - Map codes to be stored in a field named Map_code

Note: An initial Personal Geodatabase will be supplied by CEH as a starting point for commencing digitization. **Note:** 5 squares have already been digitized as examples (63,68,195,414,869).

- Linear and point features are not to be digitized.

- Site location reference number (square number) is found written on the map in the following format '6/22'  where site location reference is the number **after** the slash. E.g. If map is labelled 6/22, the site number is 22. **Ignore the 1st number**. The site number does not need to be stored within the data. A file named BLKDATA_78 will

be supplied with the locations of all sample squares (see below), where **BLK** is the square number.
- Digitizing log to be kept with time taken to digitize square and notes regarding any specific problems (Digitizing_log_CS1978.doc)

2.2 Topology Rules

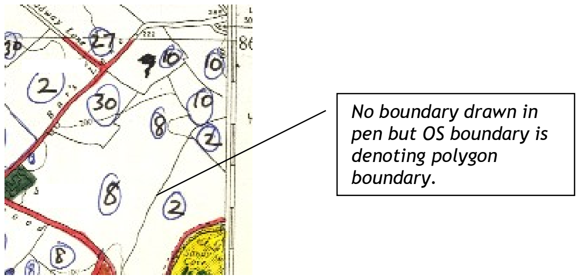
- Areas_78:
    - polygons must not overlap
    - polygons should not have gaps within each survey square
    - Areas_78 should coincide with BLKDATA_78 and vice-versa.
    - No polygons to be created having an areas less than 400m^{2}. Areas less than this on the paper map should be subsumed into the adjacent largest polygon.

2.3 Specific Issues Relating to Map Coding

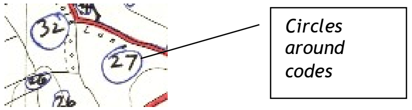
- The values for populating the attribute 'Map_Code' in Areas_78 polygons must be valid codes i.e. must appear on code list (Annex 1 / *Codes_1978.xls*)

- Digitize contiguous area with the same map code as one area. Where black boundary lines have been drawn, follow these as polygon boundaries.

Where there are no hand drawn boundaries separating areas, follow OS boundary lines (only do this where there is an obvious drawn boundary missing between 2 heterogeneous boundaries).

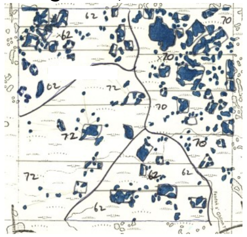


- Ignore circles around map codes unless they are obviously supposed to denote an area.



- Areas that have no obvious label, and a label cannot be determined (E.g. Sea, Lake, Railway, Recreation ground etc.) use code 22 (unassigned).

- Areas less than 400m^2 (minimum mappable unit: MMU) not to be digitized but subsumed by adjacent largest polygon (ignore Map Code for the smaller polygon). **An exception to this would be if the polygon falls on the edge of a square.**
  Ponds, for example in square 1162 (see below) may be a issue. Ponds above the MMU should be digitised, those below should be ignored.



- Ignore codes on the map preceded by an X e.g. X1, X2, X3, X4, X5. These are not relevant to the area codes.

- If there appears to be 2 codes for one area e.g. 2*3, use the first code

- Any areas which has been split by a road or stream, thereby making it unclear which code should be assigned to the polygon, should be given a code of **999**, in order that CEH can look at it later on, and make a decision on the final code allocation

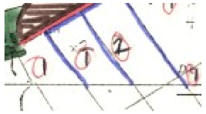
- Any comments on the map, or apparent changes to codes should be added to a column in the dataset called ‘Comments’.

CEH C03689 Countryside Survey 1978: Habitat Map Digitization Project
Digitizing Protocol

Notes on colour codes:

Rivers/Streams (Blue)

- Blue lines measuring 1mm wide or less on the paper map should not be digitized.



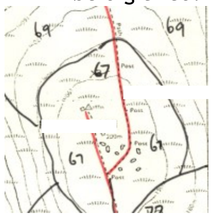
Not to be digitized

- Blue lines greater than 1mm should be digitized as an area and coded 41 (Canal/stream) unless the area created would constitute an area smaller than the MMU.

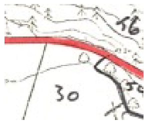
To be digitized as area

Roads (Red)

- Red lines measuring 1mm wide or less on the paper map should not be digitized.



- Thick red lines greater than 1mm should be digitized as an area and coded 36 (Road) unless the area created would constitute an area smaller than the MMU.

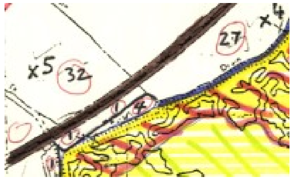


- Ignore red lines within urban areas (part of urban polygon).

Railways (Brown)

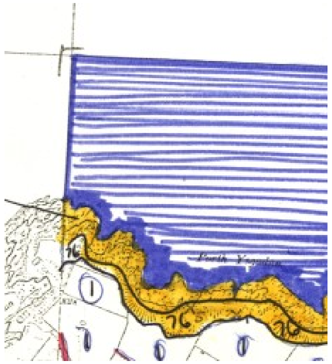
- Brown lines measuring 1mm wide or less on the paper map should not be digitized.

- Thick brown lines greater than 1mm should be digitized as an area and coded 39 (railway) unless otherwise labelled, and unless the area created would constitute an area smaller than the MMU.
-



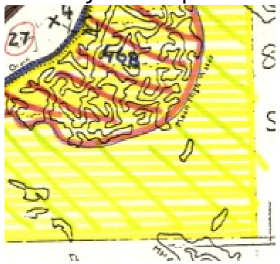
Water bodies (Blue)
Where a lake/water body is coloured blue, digitize as an area and code as 42 (Lake) unless the area created would constitute an area smaller than the MMU.

Sea (Blue)
Digitize areas of sea at low water mark (may be coloured blue). Code as 81 (Sea) – Ignore any other given codes for this.



Inter-tidal Zones (Yellow)
Digitize yellow inter-tidal zones (see above) as 40 (Cliffs/sand/mud) unless labelled otherwise (possibly 46 (Rock), 76 (Maritime grassland) or 79 (Salt-marsh)).

Salt-marsh (Yellow/red hatching)
Areas coded with 468 should be coded as 79 (saltmarsh) this should have been adjusted in pencil in most cases.



Urban Areas (Green)
Areas coloured in a shade of green to be labelled 37 (urban)

Recreation Grounds (Orange/Red)

Areas coloured in shades of orange/red should be labelled as 44 (Formal recreation areas)

Note: Specific issues may need to be dealt with on a case by case basis

2.4 Quality / Accuracy

Spatial:

- All linework should be within 2m of that shown on supplied paper maps (unless subject to reasonable adaptation to coincide with other features).
- Digitizing to follow centre line of linework
- Follow OS lines where no other boundary line has been drawn between polygons
- No contraventions of topology rules as specified.

Attributes:

- No duplicates of Area_78 unique_id
- No missing Map_Code values.
- No invalid Map_Code values.

3. Schedule of completion

It is anticipated that the maps will be completed by 7th August.

4. Technical Queries

In the event of technical queries please contact:
***In the first instance*: Claire Wood,** [redacted]
Or: Simon Smart, [redacted]

Digital files that will be provided:

File Description	**File name**
Code list	Codes_1978.xls
Initial Personal Geodatabase:	CS_Maps_1978.mdb
- BLKDATA_78 (locations of boundary squares for all maps) In pgdb, above.	BLKDATA_78 **BLK REFERS TO SURVEY SQUARE NUMBER**
- Areas_1978 (includes 5 squares already digitized). In pgdb, above.	Areas_1978
Digitizing log template	Digitizing_log_CS1978.doc
Copy of this protocol document	CS1978_Digitizing_Protocol_JUN09.doc
JPG versions of survey squares	CS1978_Map X.pdf **NOTE: NUMBER OF JPG DOES NOT REFER TO SURVEY SQUARE NUMBER**

Above to be supplied on CD.

Final product: Feature class named Areas_78 (already set up in pgdb above), containing all polygons for all squares provided.

Appendix iv: Examples of Supporting Metadata Documents

Dataset Documentation

<u>Scottish Pinewoods Survey 1971</u> (Native Pinewood Survey)

Document version 1.1 4/9/2015

Prepared by C.M. Wood[1], D. Caffrey[2] & R.G.H. Bunce[2].
[1]CEH Lancaster, Library Avenue, Bailrigg, Lancaster. LA1 4AP.
[2] Formerly of the Institute of Terrestrial Ecology, Merlewood, Grange-over-Sands, Cumbria.

Contents

Dataset Series Name: **Scottish Pinewood Survey 1971**

Dataset Description: A detailed ecological survey of the Scots Pine woodland habitats within Scotland. In all, 27 woods from throughout northern Scotland were identified as the major remaining native pinewoods, and within each wood 16 randomly selected 200m^2 plots were surveyed (26 of the woods were surveyed in 1971, with 1 extra wood surveyed in 1972). Details about the trees, ground flora, soil, habitat types as well as general plot information were collected for each plot using standardized procedures and coding systems.

Geographic Coverage: Scotland

Time Period: 1971-72

Data Categories:

Vegetation Data:	Vascular plants. Bryophytes. Trees, saplings & shrubs.
Soil Data:	Horizon depths and descriptions. pH.
Habitat Data:	Habitat categories. Slope. Aspect.

Survey Design & Methods: Samples of all known major pinewoods in Scotland.
Bunce & Shaw's 1973 standardized survey methods.

Related Datasets:

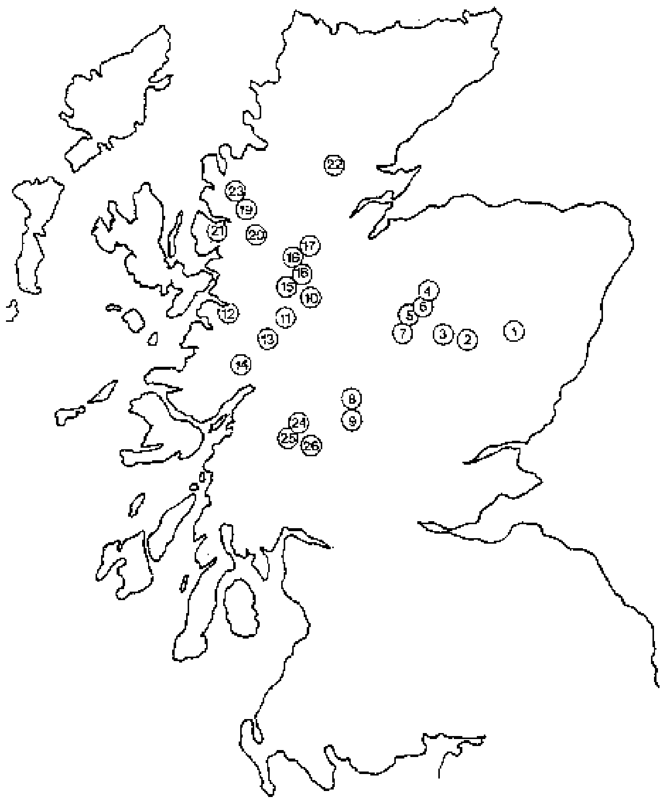
- **National Woodland Survey 1971** - carried out in the same year, using the same methodology (and repeated in 2000-2003).
  Scottish Pinewood Survey 1973 - a follow-up survey to a subset of the woods.

Key documents & publications:

- Steven H.M. & Carlisle A. (1959). *The Native Pinewoods of Scotland.* Edinburgh: Oliver & Boyd.
- Bunce R.G.H & Shaw M.W. (1971). *National Woodland Classification 1971: Handbook of Field Methods.* Unpublished document, ITE Merlewood.
- Bunce R.G.H. & Shaw M.W. (1973). A Standardized Procedure for Ecological Survey. *Journal of Environmental Management*, Vol. 1, 239-258.
- Hill M.O., Bunce R.G.H, & Shaw M.W. (1975). Indicator species analysis: a divisive polythetic method of classification and its application to a survey of native pinewoods in Scotland. *Journal of Ecology*, Vol. 63, 597-613.
- Bunce R.G.H. (1977). The range of variation in the pinewoods. In: *Native pinewoods of Scotland*, edited by R.G.H. Bunce and J.N.R. Jeffers, 10-25. Cambridge: Institute of Terrestrial Ecology.
- Goodier R. & Bunce R.G.H. (1977). The Native Pinewoods of Scotland: The current state of the resource. In: *Native pinewoods of Scotland*, edited by R.G.H. Bunce and J.N.R. Jeffers, 78-87. Cambridge: Institute of Terrestrial Ecology.

Originator Details: R.G.H. Bunce, Woodlands Research Section, Nature Conservancy - Merlewood.

1. Geographical Coverage



1 Glentanar	8 Black Wood of Rannoch	14 Ardgour	21 Shieldaig
2 Ballochbuie	9 Old Wood of Meggernie, Glen Lyon	15 Glen Affric	22 Amat
3 Mar		16 Glen Cannich	23 Loch Maree
4 Abernethy	10 Glen Moriston	17 Glen Strathfarrar	24 Black Mount
5 Rothiemurchus	11 Glengarry	18 Guisachan and Cougie	25 Glen Orchy
6 Glenmore	12 Barrisdale	19 Coulin	26 Tyndrum
7 Glen Feshie	13 Loch Arkaig and Glen Mallie	20 Achnashellach	

N.B. The Dulnan, surveyed in 1972 is located to the west of Abernethy (4) and north of Glen Feshie (7).

2. Scottish Pinewood Survey 1971 - *Overview of datasets*

2.1 Vegetation Data

In each of the woodland plots, 3 categories of vegetation data were recorded:-
i) Trees, saplings & shrubs
ii) Vascular plants
iii) Bryophytes growing on the ground

Method
A nested quadrat system was used to record the vegetation present within each of the 200m^2 plots surveyed. Individual trees were recorded throughout the plot, whilst individual saplings and shrubs were recorded in opposing quarters of the plot only. For vascular plants, species lists were recorded by 4m^2, 25m^2, 50m^2, 100m^2 & 200m^2 nested quadrats, with only previously unfound species listed as the quadrat size increases. An estimate of the % cover for each of the species throughout the plot overall was also recorded, by class size (5% bands, 1%, few/ Presence '+' has been replaced by 0.5). Samples of bryophytes were collected for later identification and species lists drawn up for the plot overall.

Parameters (refer to field handbook (Shaw and Bunce, 1971) for full descriptions)

Trees, saplings & shrubs - by individual
Species name
Diameter at breast height
Tree dead indicator
Height of widest tree in plot
Grouping of stems in a coppice stool

Vascular plants - by species and quadrat
Species name
Species name code
% cover estimate

Bryophytes - by species and plot
Species name
Species name code

Other Cover/ Abundance Data - by plot
% Litter
% Wood
% Rock
% Bare ground
% Water
% Bryophytes

2.2 Soil Data

The soil of each woodland plot was classified by horizon using the set of standard categories outlined below. pH was also measured for soil samples from each plot (top 0-15cm).

Methods
In the centre of each plot a shallow pit was dug to enable examination of the surface layers of soil, and auger samples were taken to classify lower horizons. Precise definitions for each of the descriptive categories were used and are detailed in the field handbook (Shaw and Bunce, 1971). A sample from the top 10cm was taken away for the pH analysis.

Parameters
pH
Horizon depths
Horizon descriptions (see table below)
Rocks & stone type and percentages (see table below)

Soil Horizon Descriptive Categories & Classes (refer to field handbook for full descriptions)

Litter Layer

Composition	
Tree leaves	Fern
Needles	Ericoid
Grass	Bryophyte
Herb	Wood

Organic Layer

Texture	*Moisture*
Fibrous	Very Wet
Granular	Wet
Amorphous	Damp
	Dry

Mixed Mineral / Organic Layer

Transition with mineral soil	*Colour*	*Texture*	*Moisture*	*Structure*
Sharp	Black	Clay	Very Wet	Powder
Gradual	Brown	Silt	Wet	Crumb
	Red	Sandy	Damp	Clod
	Mottled	Stony	Dry	

Leached or Eluviated Layer

Colour	*Texture*
Whitish	Clay
Greyish	Silt
	Sandy
	Stony

Weathered Mineral Layer

Deposition layer		*Colour excl dep. layer*	*Texture*	*Moisture*	*Structure*
		Yellow	Clay	Very Wet	Powder
Black	Comp.	Yellow/Brown	Silt	Wet	Crumb
Red/Brown	Uncomp.	Brown	Sandy	Damp	Clod
		Red	Stony	Dry	
		Mottled			

Underlying Material

Texture	
Clay	Stony
Silt	Rock (frag)
Sandy	Rock (solid)

Rocks & Stones in Soil

Composition (%)		*Shape (%)*	*Size range (%)*
Slate/shale	Limestone	Round	<5 cm
Sandstone	Flint	Sub-angular	5-10cm
Grit	Granite	Angular	10-20cm
Chalk	Others		>20cm

2.3 Habitat Data

Habitat types of each of the woodland plots were classified using the specific pre-defined categories given below. The slope and aspect of each plot were also measured. Additional descriptions were recorded for the whole site.

Methods
All the classes within each of the habitat categories listed which applied to the 200m^2 plot were crossed on the data sheet. Precise definitions for each category and its classes were used, and are detailed in the field handbook (Shaw and Bunce, 1971).
Slope was measured using a clinometer from the highest to lowest point in the plot, passing through the centre of the pot. The aspect was taken bearing down the slope, measured with a *Silva* compass.

Parameters
slope (°), aspect (° mag), habitat types (see table below)

Habitat Categories & Classes (refer to field handbook for full descriptions)

Trees - Management

Coppice stool	Singled coppice	Recently cut coppice	Stump hard. new
Stump hard. old	Stump con. new	Stump con. old	

Trees - Regeneration

Alder	Rhododendron	Ash	Sweet chestnut	Hazel	Yew	Rowan
Birch	Other hardwood	Hawthorn	Scots pine	Oak	Beech	Wych elm
Hornbeam		Lime	Aspen	Sycamore	Holly	Other conifers

Trees - Dead (=habitats)

Fallen broken	Fallen uprooted	Log very rotten	Fallen bnh >10cm
Hollow tree	Rot hole	Stump <10cm	Stump >10cm

Trees - Epiphytes & Lianes

Bryo. base	Bryo. trunk	Bryo. branch	Lichen trunk
Lichen branch	Fern	Ivy	Macrofungi

Habitats - Rock

Stone <5cm	Rocks 5-50cm	Boulders >50cm	Scree
Rock outcrop <5m	Cliff >5m	Rock ledges	Bryo covered rock
Gully	Rock piles	Exp. gravel/sand	Exp. min. soil

Habitats - Aquatic

small pool 1m^2	Ditch/drain dry	Aquatic veg.	Pond/lake >20m^2	Stream/river slow Marsh/bog
stream/river fast	Pond 1-20m^2	Ditch/drain wet	Spring	

Habitats - Open

Gld. 5-12m	Gld. >12m	Rocky knoll <12m	Rocky knoll >12m
Path <5m	Ride >5m	Track non-prep	Track metalled

Habitats - Human

Wall dry	Wall mortared	Wall ruined	Embankment
Soil excavated	Quarry/mine	Rubbish dom.	Rubbish other

Habitats - Vegetation

Blackthorn	Bracken dense	Rose clump	Rhododendron thicket	Macfungi soil	Umbel. clump	
thicket	Leaf drift Hawthorn	Moss bank	W.herb clump	Bramble clump	Grass bank	
Nettle clump	thicket	Herb veg >1m	Fern bank		Macfungi. wood	

Animals (mainly signs of)

Sheep	Cattle	Horse/pony	Pig
Red deer	Other deer	Rabbit	Badger
Fox	Mole	Squirrel	Anthill
Corpse/bones	Spent cartridges		

3. *Scottish Pinewood Survey 1971 – Survey Design & Methods*

Original Purpose of Survey
To provide a more precise definition of the range of ecological variation in pinewoods than previously existed before 1971; to enable the development of an integrated conservation strategy for the remaining native pinewoods in Scotland.

Sample Design
All the major native pinewoods identified in the book by Steven and Carlisle (1959) were included in the 1971 survey. In addition the Dulnan, the other remaining major pinewood, was surveyed in 1972.

In each of the selected woods, 16 randomly assigned 200m^2 quadrats were selected for the survey.

Survey Methods
The methods outlined in *Bunce & Shaw (1973)* were used for this survey. Further details of exactly what and how the information was recorded, including definitions of any classifications used are given in the field handbook. Bunce R.G.H & Shaw M.W. (1971). *National Woodland Classification 1971: Handbook of Field Methods.* Unpublished document, ITE Merlewood.

Summaries of the methods used are as follows:

Vegetation Data
A nested quadrat system was used to record the vegetation present within each of the 200m^2 plots surveyed. Individual trees were recorded throughout the plot, whilst individual saplings and shrubs were recorded in opposing quarters of the plot only. For vascular plants, species lists were recorded by 4m^2, 25m^2, 50m^2, 100m^2 & 200m^2 nested quadrats, with only previously unfound species listed as the quadrat size increases. An estimate of the % cover for each of the species throughout the plot overall was also recorded, by class size (5% bands, 1%, few). Samples of bryophytes were collected for later identification and species lists drawn up for the plot overall.

Soil Data
The soil of each plot surveyed was classified by horizon using a set of standard categories. In the centre of each plot a shallow pit was dug to enable examination of the surface layers of soil, and auger samples were taken to classify lower horizons. Precise definitions for each of the descriptive categories used are detailed in the field handbook. A sample from the top 10cm was taken away for pH analysis.

Habitat Data
Habitat types of each of the woodland plots were classified using pre-defined categories and descriptive classes. All classes within each of the habitat categories listed on the data sheets which applied to the 200m^2 plot were crossed on the data sheet. Precise definitions for each habitat category and its classes were provided, and are detailed in the field handbook. Slope was measured using a clinometer from the highest to lowest point in the plot, passing through the centre of the pot. The aspect was taken bearing down the slope, measured with a *Silva* magnetic compass.

Descriptions were also recorded for each woodland site.

4. Summary of available data per site

Site no.	Approx. start date	Site name	OSGR	Site description	Plot description	Soil	Ground flora	Tree data	Plot data
1	16 July 1971	Glentanar	NO470920	Y	Y	Y	Y	Y	Y
2	29 July 1971	Ballochbuie	NO200895	Y	Y	Y	Y	Y	Y
3	18 July 1971	Mar	NO035932	Y	Y	Y	Y	Y	Y
4	12 July 1971	Abernethy	NH990180	Y	Y	Y	Y	Y	Y
5	22 July 1971	Rothiemurchus	NH920080	Y	Y	Y	Y	Y	Y
6	22 July 1971	Glenmore	NH980090	Y	Y	Y	Y	Y	Y
7	23 July 1971	Glen Feshie	NN845990	Y	Y	Y	Y	Y	Y
8	25 July 1971	Black Wood of Rannoch	NN580560	N	N	N	N	Y	Y
9	27 July 1971	Old Wood of Meggernie, Glen Lyon	NN555455	Y	Y	Y	Y	Y	Y
10	22 August 1971	Glen Moriston	NH310120	Y	Y	Y	Y	Y	Y
11	15 August 1971	Glengarry	NH230010	Y	Y	Y	Y	Y	Y
12	14 August 1971	Barisdale	NG890030	Y	Y	Y	Y	Y	Y
13	21 August 1971	Loch Arkaig and Glen Mallie	NN170875	Y	Y	Y	Y	Y	Y
14	18 August 1971	Ardgour	NM960713	Y	Y	Y	Y	Y	Y
15	13 July 1971	Glen Affric	NH145225	Y	Y	Y	Y	Y	Y
16	24 August 1971	Glen Cannich	NH160300	Y	Y	Y	Y	Y	Y
17	30 October 1971	Glen Strathfarrar	NH370390	Y	Y	Y	Y	Y	Y
18	26 August 1971	Guisachan and Cougie	NH298235	Y	Y	Y	Y	Y	Y
19	07 August 1971	Coulin	NG995557	Y	Y	Y	Y	Y	Y
20	11 August 1971	Achnashellach	NH035470	Y	Y	Y	Y	Y	Y
21	10 August 1971	Shieldaig	NG820524	Y	Y	Y	Y	Y	Y
22	03 August 1971	Amat	NH460855	Y	Y	Y	Y	Y	Y
23	06 August 1971	Loch Maree	NH010609	Y	Y	Y	Y	Y	Y
24	29 July 1971	Black Mount	NN350455	Y	Y	Y	Y	Y	Y
25	30 July 1971	Glen Orchy	NN250360	Y	Y	Y	Y	Y	Y
26	28 July 1971	Tyndrum	NN330280	Y	Y	Y	Y	Y	Y
27	05 August 1972	Dulnan	NH830180	Y	Y	Y	Y	Y	Y

Y = Yes
N= No

***Note:* Additional plots (up to 50) were undertaken at site 4, Abernethy.**

5. Data Tables and Descriptions

Scots_Pine_1971_Sites.csv

Description: Approximate locations of surveyed Scots Pine woodlands (point features).

Column Name	Type	Description
ID	Long Integer	Site ID number (1-27)
NAME	Text	Site name
OSGR	Text	OS grid reference of site point
POINT_X	Double	Easting of site in metres (OSGB1936, BNG)
POINT_Y	Double	Northing of site in metres (OSGB1936, BNG)

SCOTS_PINE_1971_SITE_INFO.csv

Description: Descriptions of surveyed sites (whole woodland), surveyed plots (individual plots), including habitat descriptions, animal descriptions and tree descriptions. Includes soil horizon descriptive categories for plots. Includes site names, plot slope and aspect, and survey dates.

Column name	Type	Description
SITE_NO	Number	Site number (1-27)
PLOT_NO	Number	Plot number (1-16, site 4; 1 - 50)
CODE	Number	Code (from recording sheet)
DESCRIPTION	Text	Description
CODE_PC	Number	Percentage applying to code (where relevant)
CODE_GROUP	Text	Code grouping (see field sheets)
CODE_GROUP_DESCRIPTION	Text	Description of code grouping
PLOT_DATE	Date	Date on which plot was surveyed
PLOT_SLOPE	Number	Slope in degrees
PLOT_ASPECT	Number	Aspect in degrees magnetic
DATA_SHEET	Text	Field sheet on which codes were recorded

SCOTS_PINE_1971_TREE_DATA.csv

Description: Tree species data from each plot, including diameter at breast height measurements (DBH).

Column	Type	Description
SITE_NO	Number	Site number (1-27)
PLOT_NO	Number	Plot number (1-16, site 4; 1 - 50)
TREE_NO	Number	Tree number (identical tree numbers denote different stems on the same

		individual tree)
NEST	Number	Nest number, 1-4 (see field handbook (Shaw and Bunce, 1971)). Denotes quarters of the plot (not the same as the ground flora nests).
TREE_TYPE	Text	Type of tree (Tree, Sapling or Shrub)
SPECIES	Text	Species of tree (Stace, 1997)
DBH	Number	Diameter at breast height, measurement in cm.
DEAD	Text	'D' if stem is dead
TREE_HT	Number	Height of widest tree in plot m (sp. of tree in 'Species' column)
DATA_SHEET	Text	Field sheet on which codes were recorded

SCOTS_PINE_1971_GROUND_FLORA.csv

Description: Ground flora records (vascular plants and bryophytes) for Scots Pine 1971 plots, including bare ground, total bryophytes, litter, rock and wood cover. Nomenclature follows (Stace, 1997).

Column name	Type	Description
SITE_NO	Number	Site number (1-27)
PLOT_NO	Number	Plot number (1-16, site 4; 1 - 50)
NEST	Number	Nest number 1-5 (see field handbook (Shaw and Bunce, 1971))
COVER	Number	% cover in nest (0.5 denotes 'present')
BRC_NUMBER	Number	Biological Record Centre species number, where available
BRC_NAME	Text	Scientific name or description
COMMON_NAME	Text	Common name where available
GROWTH_FORM	Text	Growth form of species. *aq - Aquatic, b - Bryophyte, f - Forbs, fe - Ferns, g - Grass, l - Lichen, ma - Marine alga, m - Other monocots, s - Sedge, ss - Dwarf shrub, w - Woody*

SCOTS_PINE_1971_SOIL_DATA.csv

Description: Soil data, per plot. Soil pH and horizon depths.

Column name	Type	Description
SITE_NO	Number	Site number (1-27)
PLOT_NO	Number	Plot number (1-16, site 4; 1 - 50)
FLORA_RECORDER	Text	Initials of flora surveyor
SOIL_PH	Number	Soil pH value
LITTER_FROM	Number	Depth of litter layer in cm from
LITTER_TO	Number	Depth of litter layer in cm to
ORGANIC_FROM	Number	Depth of organic layer in cm from
ORGANIC_TO	Number	Depth of organic layer in cm to
MIXED_FROM	Number	Depth of mixed layer in cm from
MIXED_TO	Number	Depth of mixed layer in cm to
LEACHED_FROM	Number	Depth of leached layer in cm from
LEACHED_TO	Number	Depth of leached layer in cm to
WEATHER_FROM	Number	Depth of weathered layer in cm from
WEATHER_TO	Number	Depth of weathered layer in cm to
UNDER_DEPTH	Number	Depth of underlying material from
DATA_SHEET	Text	Field sheet on which codes were recorded

6. References

Shaw, M. W. & Bunce, R. G. H. (1971) National Woodlands Classification 1971 Handbook of Field Methods. Merlewood Research Station.
Stace, C. (1997) *New flora of the British Isles.* Cambridge University Press, Cambridge.
Steven, H. M. & Carlisle, A. (1959) *The native pinewoods of Scotland.* Oliver & Boyd, Edinburgh.

7. Acknowledgements

- *Survey management: Bob Bunce, Wally Shaw*
- *Field surveyors: C. Eccles, D.J. Taylor, P. Bassett, K. H. Chorlton, K. Wilson, M. Ball, M. J. Bottomly, M. W. Shaw, R. Bunce, D.J. Taylor, P. Bassett*
- *Data management and documentation: Claire Wood, Caroline Hallam, Deirdre Caffrey*

Ecological survey of 'Key Habitats' in England, 1992-93

Dataset Documentation

Dataset Name: ***Habitat and vegetation data from an ecological survey of terrestrial key habitats in England, 1992-1993***

Dataset Background & Description: Since 1978, a series of national surveys (Countryside Surveys) have been carried out by the Institute of Terrestrial Ecology/Centre for Ecology and Hydrology to gather data on the natural environment. The sampling framework for these surveys is not optimised to yield data on rarer or more specialised habitats. This Key Habitat survey was commissioned in the early 1990s to carry out additional survey work into habitats which were perceived to be under threat or which represented areas of concern to the Department for the Environment. The habitats were: Lowland heath, Chalk and limestone (calcareous) grasslands, Coasts, Uplands. For reporting purposes, the survey was supplemented by additional data, particularly for river valleys and waterside landscapes (lowlands), from Countryside Survey 1990.

Geographic Coverage: England

Time Period: 1992-93

Data Categories:

Vegetation Data:	Vascular plants. Bryophytes. Lichens.	
Boundary data:	Boundary descriptions at grid points across each 1km survey square	
Habitat Data:	Land cover data at grid points across each 1km survey square	

Survey Design & Methods: Stratified sampling across entire area. Bunce & Shaw's 1973 standardized survey methods with modifications – refer to field handbooks.

Related Datasets: Countryside Survey http://www.countrysidesurvey.org.uk/

Key documents & publications:

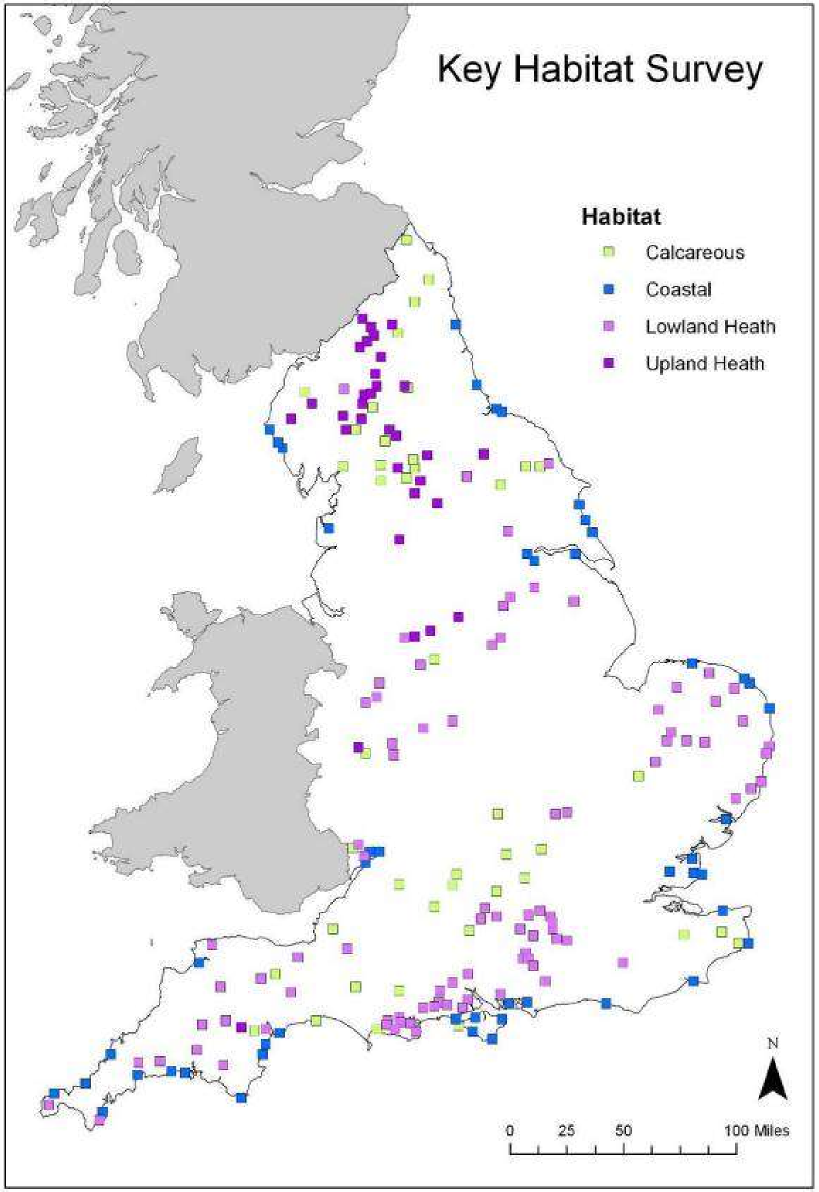
- Barr, C.J. (1993) *Changes in Key Habitat: Surveys of Chalk & Limestone Grassland, Coastal, and Upland Landscapes, Field Handbook. Draft 3.* Grange-over-Sands, NERC/Institute of Terrestrial Ecology.
- Barr, C.J. (1992) *Changes in Key Habitat: A Survey of Lowland Heath, Field Survey Handbook.* ITE Merlewood.
- Barr, C.J. (1997) *Current Status and Prospects for Key Habitats in England, Part 6 Summary Report.* Institute of Terrestrial Ecology, University of Sheffield - Unit of Comparative Plant Ecology, Environmental Resource Management Ltd, Department of the Environment, Transport and the Regions.
- Bunce R.G.H & Shaw M.W. (1973). A Standardized Procedure for Ecological Survey. *Journal of Environmental Management*, Vol. 1, 239-258.

Originator Details: *Survey & Project Management:* Barr C.J., Bunce, R.G.H., Cummins, R.P., Hallam, C.J. Hornung, M.: Institute of Terrestrial Ecology, Merlewood, Grange over Sands, Cumbria.
Data management: Wood, C.M.: Centre for Ecology & Hydrology, Lancaster.

Field surveyors: *H. Adams, T. Barden, E. Biron, R. Cummins, J. Davis, J. Day, D. De La Pole, C. Hallam, M. Harrison, R. Hewison, J. Hobbs, C. Kanefsky, G. Levine, M. Marler, D. McCutcheon, L. McDonnell, M. Ness, K. Pollock, S. Walters, M. Webb*

Quality control: *All data were collected by trained surveyors. Collected data have been checked and verified as far as possible.*

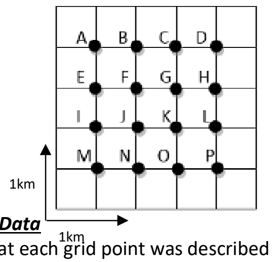
Geographical Coverage



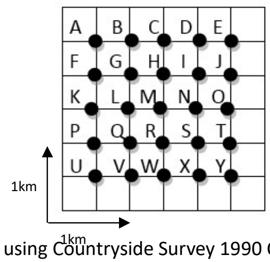
Key Habitat Survey of England 1992-93 - Habitat Data & Boundary Data

At each point on a regular grid (16 or 25 depending on landscape/habitat type), within 1 km survey squares, codes describing land cover and the nearest boundary were recorded. Refer to field handbooks for more detail (Barr, 1992;Barr, 1993).

16 point grid



25 point grid



Land Cover Data

Land Cover at each grid point was described using Countryside Survey 1990 Codes. A list of these codes can be found in Appendix II. The codes reflected the "mappable unit" in which the point fell. The minimum mappable unit was 400m². Each mappable unit was determined by the constancy of the codes which described it. If one characteristic (e.g. cover of a dominant plant species) was sufficiently different from an adjacent area to be given a different code then a new mappable unit was recognised.

Boundary data

The nearest vertical boundary to each grid point (within 100m) was described using Countryside Survey 1990 Codes. A list of these codes can be found in Appendix III. The point on the boundary which was nearest to the grid point was recorded as part of a length which can be coded constantly as part of a single unit of not less than 20m (the minimum mappable length). If the nearest point on the boundary was part of a longer length, then the coding reflected the variability of the longer length. Vertical boundaries included: hedgerows, walls, fences, earth/stone banks acting as field boundaries; any combination of these.

Key Habitat Survey of England 1992-93 - Vegetation Data

Detailed plant species information was collected from a series of plots (as outlined in table 1). In this survey, 4 types of plot were used:

a) Main (X) plots: these were either $4m^2$ or $200m^2$ (nested), depending on the landscape type (see table 1) and were located at five of the grid points within the square ($200m^2$) or all grid points ($4m^2$).
b) Targeted (Y) plots: these are $4m^2$ and 5 were placed in semi-natural vegetation types/ habitats that had not been covered by the main plots.
c) Stream side (S) and waterside (W) plots: these are 10 x 1m plots and up to 5 were placed immediately adjacent to watercourse where present, in Upland Landscapes only.
d) Roadside (R) and Verge (V) plots: these are 10x1m plots and up to 5 were placed immediately adjacent to roads where present in Limestone and Chalk (calcareous) landscapes only.

Table 1: Summary of Survey Sites *(refer to field handbooks)*

Habitat	No. of 1km squares	Map Grid	X Plots ($200m^2$)	X plots (2x2m)	Y Plots (2x2m)	SW plots (10x1m)	RV plots (10x1m)	Year surveyed
Calcareous	43	16 points, A-P	-	5 plots recorded at AJGDP	5 at locations selected by surveyor	-	5 plots adjacent to roadsides.	1993
Coastal	49	25 points, A-Y	5 plots recorded at points ALITW	-	5 at locations selected by surveyor	-	-	1993
Lowland Heath	89	25 points, A-Y	-	25 plots, on grid.	-	-	-	1992
Upland Heath	32	16 points, A-P	5 plots recorded at AJGDP	-	5 at locations selected by surveyor	5 plots adjacent to watercours es	-	1993
Total	*213*	*-*	*240*	*675*	*620*	*150*	*201*	*-*

Table 2: Summary of Vegetation Plots

Habitat	No. of 1km squares	X Plots (200m2)	X plots (4m2)	Y Plots (2x2m)	SW plots (10x1m)	RV plots (10x1m)
Calcareous	43	-	122	215	-	81(R) 120(V)
Coastal	49	92	-	245	-	-
Lowland Heath	89	-	553	-	-	-
Upland Heath	32	148	-	160	60 (S) 90 (W)	-
Total	*213*	*240*	*675*	*620*	*150*	*201*

Note: 35 squares have no plot data, mostly in the lowland heath habitat (not recorded due to being located in arable or intensively improved grassland)

Key Habitat Survey of England 1992-93 - Datasets

List of tables

Table Name	Description
Key_Habitat_Survey_Squares.csv	List of set of 1km squares surveyed, for each habitat type
Key_Habitat_Plot_Headers.csv	Vegetation plot header information and location
Key_Habitat_Plot_Species.csv	List of plant species recorded within each vegetation plot
Key_Habitat_LC_Features.csv	Land cover data mapped at each grid point within squares
Key_Habitat_LC_Species.csv	Species details pertaining to land cover features in Key_Habitat_LC_Features table
Key_Habitat_BD_Features.csv	Boundary data mapped at each grid point within squares
Key_Habitat_BD_Species.csv	Species details pertaining to boundary features in Key_Habitat_BD_Features table

Table Relationships

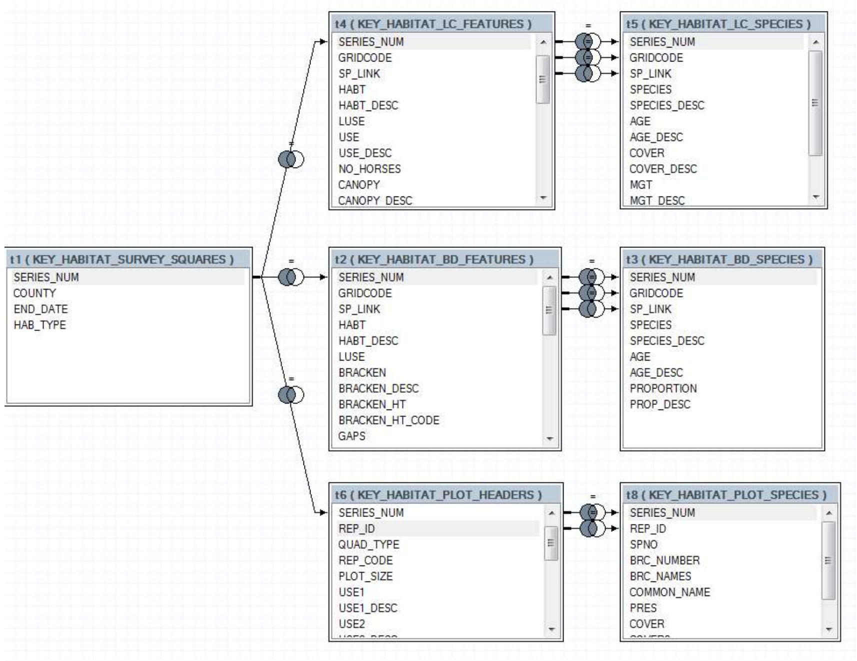


Table Descriptions

Key_Habitat_Survey_Squares.csv

Information from front cover of Field Assessment Booklet regarding 1km survey square. Site code, county in which located, end of survey date, habitat type.

Column name	Description	Type
SERIES_NUM	Survey square number	Text
COUNTY	County square located in	Text
END_DATE	Date survey square finished	Date
HAB_TYPE	Key Habitat (landscape) type	Text

Key_Habitat_Plot_Headers.csv

Vegetation plot header information and location (where known).

Column name	Description	Type
HAB_TYPE	Key Habitat (landscape) type	Text
SERIES_NUM	Survey square number	Text
REP_ID	Code of plot	Text
QUAD_TYPE	Type of plot: Y (targeted), X (main plot), R (roadside),V (verge), S (stream),W (waterside) *(see handbook)*	Text
REP_CODE	Number of plot if X, Y, R or V plot types (rather than alphabetic grid codes)	Text
PLOT_SIZE	Size of plot: Y= 2x2m, X= 200m^2 (14.1x14.1m) **or** 2x2m, R=10x1m S=10x1m *(see handbook)*	Text
USE1	Land Use code 1 (from plot sheet)	Numeric
USE1_DESC	Land Use code 1 description	Text
USE2	Land Use code 2 (from plot sheet)	Numeric
USE2_DESC	Land Use code 2 description	Text
USE3	Land Use code 3 (from plot sheet)	Numeric
USE3_DESC	Land Use code 3 description	Text
PHYS	Physiology code (from plot sheet)	Numeric
PHYS_DESC	Physiology description	Text
SLOPE	Slope description. (from plot sheet) Flat (no slope discernable), Slight (<5° by eye), Moderate (6-15°), Steep (>15)	Text
ASPECT	Aspect in degrees from north. None if flat (from plot sheet)	Numeric
SHADE	Shade description (from plot sheet). Full/open/partial/none	Text
GRAZING	Grazing description (from plot sheet)	Text
SUBSTRATE	Substrate description (from plot sheet)	Text
SOIL_DEPTH	Soil depth (from plot sheet)	Text
DESCRIPTION1	Site description 1 – notes (from plot sheet)	Text
DESCRIPTION2	Site description 2 – notes (from plot map sheet)	Text

Key_Habitats_Plot_Species.csv

List of plant species and covers recorded within each vegetation plot. *Note: Bare ground and total bryophytes not entered in R/V or S/W plots.*

Column name	Description	Type
SERIES_NUM	Survey square number	Text
REP_ID	Plot code description, sometimes the same as Gridcode in squares where plots were done on the 16/25 point grid	Text
SPNO	Species code number	Numeric
BRC_NUMBER	Biological record centre species code	Text
BRC_NAMES	Biological record centre species name *(following New flora of the British Isles, Stace, C., Cambridge University Press, 2010)*	Text
COMMON_NAME	Common name of species where available	Text
PRES	Quadrat nest number in which species first appears *(see handbook).* If relevant COVER or COVER2 is 0 and PRES >=1, then species is present at <5% cover. For Total Bryophytes and Lichens, 0 = whole quadrat.	Numeric
COVER	Cover % for inner quadrat 4m^2 where plots are 200m^2, otherwise whole plot for un-nested plots	Numeric
COVER2	Cover % for entire plot 200m^2 where plots are 200m^2, otherwise N/A.	Numeric
HAB_TYPE	Key Habitat (landscape) type	Text

***Key_Habitat_LC_Features*.csv**

Land cover data mapped at each grid point within survey squares.

Column Name	Description	Type
HAB_TYPE	Key Habitat (landscape) type	Text
SERIES_NUM	Survey square number	Text
GRIDCODE	Location of plot on square grid	Text
SP_LINK	Additional link to *Key_Habitat_LC_Species.csv* table	
HABT	Habitat type code	Numeric
HABT_DESC	Habitat type description	Text
LUSE	Land Use Code *(see appendix V)*	Text
USE	Use code	Numeric
USE_DESC	Use description	Text
NO_HORSES	Number of horses	Numeric
CANOPY	Canopy code	Numeric
CANOPY_DESC	Canopy description	Text
DESC	Description code	Numeric
DESC_DESC	Description of description code	Text
FOF	Forestry feature type code	Numeric
FOF_DESC	Forestry feature description	Text
FORBS	Forbs cover code	Numeric
FORBS_DESC	Forbs cover description	Text
HEATHER	Heather type code	Numeric
HEATHER_DESC	Heather type description	Text
OTHER	Additional other code	Numeric
OTHER_DESC	Additional other description	Text

PEAT	Peat code	Numeric
PEAT_DESC	Peat cover description	Text
OTHER_USE	Additional use code	Numeric
OTHER_USE_DESC	Additional use description	Text
WEEDPROP	Proportion of weeds code	Numeric
WEEDPROP_DES	Proportion of weeds code	Text

Key_Habitat_LC_Species.csv

Species details pertaining to land cover features in Key_Habitat_LC_Features table

Column Name	Description	Type
SERIES_NUM	Survey square number	Text
GRIDCODE	Location of plot on square grid	Text
SPECIES	Species code	Numeric
SP_LINK	Additional link to *Key_Habitat_LC_Features.csv* table	Numeric
SPECIES_DESC	Description of species	Text
AGE	Age code	Numeric
AGE_DESC	Description of age of species	Text
COVER	Percentage cover code of species	Numeric
COVER_DESC	Description of Percentage cover of species. *Note: sometimes this refers to the cover of the overall feature - these should be cover codes (COVER) 651-655 but occasionally these have been used for species and there is some confusion.*	Text
MGT	Management code	Numeric
MGT_DESC	Description of management	Text
HEIGHT	Height of species	Numeric
HEIGHT_DESC	Description of height	Text

Key_Habitat_BD_Features

Boundary data mapped at each boundary nearest to grid points within survey squares.

Column Name	Description	Type
SERIES_NUM	Survey square number	Text
GRIDCODE	Location of plot on square grid	Text
SP_LINK	Additional link to *Key_Habitat_BD_Species.csv* table	Numeric
HABT	Habitat/feature code	Numeric
HABT_DESC	Habitat/feature description	Text
LUSE	Land Use code *(see appendix V)*	Text
BRACKEN	Bracken code	Numeric
BRACKEN_DESC	Bracken description	Text
BRACKEN_HT_CODE	Bracken height code	Numeric
BRACKEN_HT	Bracken height description	Text
GAPS	Gaps code	Numeric
GAPS_DESC	Gaps description	Text

MGT	Management code	Numeric
MGT_DESC	Management description	Text
STOCK	Stock proof code	Numeric
STOCK_DESC	Stock proof description	Text
USE	Use code	Numeric
USE_DESC	Use description	Text
DESC1	Description 1 code	Numeric
DESC1_DESC	Description 1 description	Text
DESC2	Description 2 code	Numeric
DESC2_DESC	Description 1 description	Text
DESC3	Description 3 code	Numeric
DESC3_DESC	Description 3 description	Text
OTHER	Other information	Numeric
OTHER_DESC	Other information description	Text
CANOPY	Canopy code	Numeric
CANOPY_DESC	Canopy description	Text
HEIGHT	Feature height code	Numeric
HEIGHT_DESC	Feature height description	Text
HAB_TYPE	Key Habitat type	Text

Key_Habitat_BD_Species

Species details pertaining to land cover features in Key_Habitat_BD_Features table

Column Name	Description	Type
SERIES_NUM	Survey square number	Text
GRIDCODE	Location of plot on square grid	Text
SP_LINK	Link to *Key_Habitat_BD_Features.csv* table	
SPECIES	Species code	Long integer
SPECIES_DESC	Species description	Text
AGE	Age code (of trees)	Long Integer
AGE_DESC	Age description (of trees)	Text
PROPORTION	Proportion of cover of species code	Long Integer
PROP_DESC	Proportion of cover of species description	Text

Appendix 1 - List of Square codes and habitats/landscape types

Habitat	Square code	Habitat	Square code	Habitat	Square code	Habitat	Square code	Habitat	Square code
Calcareous	**AD1**	Lowland Heath	**A1**	Lowland Heath	**D7**	Upland Heath	**MD1**	Coastal	**ED10**
Calcareous	**AD10**	Lowland Heath	**A10**	Lowland Heath	**D8**	Upland Heath	**MD10**	Coastal	**ED2**
Calcareous	**AD101**	Lowland Heath	**A11**	Lowland Heath	**E1**	Upland Heath	**MD2**	Coastal	**ED3**
Calcareous	**AD11**	Lowland Heath	**A12**	Lowland Heath	**E10**	Upland Heath	**MD3**	Coastal	**ED5**
Calcareous	**AD13**	Lowland Heath	**A2**	Lowland Heath	**E11**	Upland Heath	**MD4**	Coastal	**ED6**
Calcareous	**AD2**	Lowland Heath	**A3**	Lowland Heath	**E12**	Upland Heath	**MD5**	Coastal	**ED7**
Calcareous	**AD4**	Lowland Heath	**A4**	Lowland Heath	**E2**	Upland Heath	**MD7**	Coastal	**ED8**
Calcareous	**AD5**	Lowland Heath	**A5**	Lowland Heath	**E3**	Upland Heath	**MD8**	Coastal	**ED9**
Calcareous	**AD6**	Lowland Heath	**A6**	Lowland Heath	**E4**	Upland Heath	**MD9**	Coastal	**EN10**
Calcareous	**AD7**	Lowland Heath	**A7**	Lowland Heath	**E5**	Upland Heath	**MN1**	Coastal	**EN2**
Calcareous	**AD8**	Lowland Heath	**A8**	Lowland Heath	**E6**	Upland Heath	**MN2**	Coastal	**EN201**
Calcareous	**AN1**	Lowland Heath	**A9**	Lowland Heath	**E7**	Upland Heath	**MN3**	Coastal	**EN3**
Calcareous	**AN10**	Lowland Heath	**B1**	Lowland Heath	**E8**	Upland Heath	**MN4**	Coastal	**EN4**
Calcareous	**AN2**	Lowland Heath	**B2**	Lowland Heath	**E9**	Upland Heath	**MN5**	Coastal	**EN6**
Calcareous	**AN4**	Lowland Heath	**B3**	Lowland Heath	**F1**	Upland Heath	**MN6**	Coastal	**EN7**
Calcareous	**AN5**	Lowland Heath	**B4**	Lowland Heath	**F2**	Upland Heath	**MN9**	Coastal	**EN9**
Calcareous	**AN6**	Lowland Heath	**B5**	Lowland Heath	**F3**	Upland Heath	**TD10**	Coastal	**HD1**
Calcareous	**AN7**	Lowland Heath	**B6**	Lowland Heath	**F4**	Upland Heath	**TD101**	Coastal	**HD10**
Calcareous	**AN8**	Lowland Heath	**B7**	Lowland Heath	**F5**	Upland Heath	**TD11**	Coastal	**HD101**
Calcareous	**AN9**	Lowland Heath	**B8**	Lowland Heath	**F6**	Upland Heath	**TD12**	Coastal	**HD3**
Calcareous	**BD1**	Lowland Heath	**C1**	Lowland Heath	**F7**	Upland Heath	**TD13**	Coastal	**HD4**
Calcareous	**BD11**	Lowland Heath	**C10**	Lowland Heath	**F8**	Upland Heath	**TD14**	Coastal	**HD5**
Calcareous	**BD13**	Lowland Heath	**C11**	Lowland Heath	**G1**	Upland Heath	**TD15**	Coastal	**HD6**
Calcareous	**BD14**	Lowland Heath	**C12**	Lowland Heath	**G10**	Upland Heath	**TD2**	Coastal	**HD7**
Calcareous	**BD15**	Lowland Heath	**(L)C13**	Lowland Heath	**G11**	Upland Heath	**TD4**	Coastal	**HD9**
Calcareous	**BD16**	Lowland Heath	**(L)C14**	Lowland Heath	**G12**	Upland Heath	**TD6**	Coastal	**HN1**
Calcareous	**BD17**	Lowland Heath	**(L)C15**	Lowland Heath	**(L)G13**	Upland Heath	**TD7**	Coastal	**HN10**
Calcareous	**BD18**	Lowland Heath	**(L)C16**	Lowland Heath	**(L)G14**	Upland Heath	**TD8**	Coastal	**HN2**
Calcareous	**BD19**	Lowland Heath	**(L)C17**	Lowland Heath	**(L)G15**	Upland Heath	**TD9**	Coastal	**HN3**
Calcareous	**BD21**	Lowland Heath	**C2**	Lowland Heath	**(L)G16**	Upland Heath	**TN2**	Coastal	**HN4**
Calcareous	**BD22**	Lowland Heath	**C3**	Lowland Heath	**G2**	Upland Heath	**TN3**	Coastal	**HN5**
Calcareous	**BD4**	Lowland Heath	**C4**	Lowland Heath	**G3**	Upland Heath	**TN4**	Coastal	**HN6**
Calcareous	**BD7**	Lowland Heath	**C5**	Lowland Heath	**G4**			Coastal	**HN9**
Calcareous	**BD8**	Lowland Heath	**C6**	Lowland Heath	**G5**	Lowland Heath	**H6**	Coastal	**SD1**
Calcareous	**BD9**	Lowland Heath	**C7**	Lowland Heath	**G6**	Lowland Heath	**H7**	Coastal	**SD10**
Calcareous	**BN10**	Lowland Heath	**C8**	Lowland Heath	**G7**	Lowland Heath	**H8**	Coastal	**SD2**
Calcareous	**BN101**	Lowland Heath	**C9**	Lowland Heath	**G8**			Coastal	**SD4**
Calcareous	**BN2**	Lowland Heath	**D1**	Lowland Heath	**G9**	Coastal	**SN2**	Coastal	**SD5**
Calcareous	**BN5**	Lowland Heath	**D2**	Lowland Heath	**H1**	Coastal	**SN3**	Coastal	**SD7**
Calcareous	**BN6**	Lowland Heath	**D3**	Lowland Heath	**H2**	Coastal	**SN5**	Coastal	**SD8**
Calcareous	**BN7**	Lowland Heath	**D4**	Lowland Heath	**H3**	Coastal	**SN6**	Coastal	**SD9**
Calcareous	**BN8**	Lowland Heath	**D5**	Lowland Heath	**H4**	Coastal	**SN7**	Coastal	**SN1**
Calcareous	**BN9**	Lowland Heath	**D6**	Lowland Heath	**H5**	Coastal	**SN8**	Coastal	**SN10**

Appendix II - Land Cover Codes

CODE	DESCRIPTION	
1	Cliff > 30m high	P
2	Cliff 5-30m high	P
3	Rock outcrop & cliff < 5m	P
4	Scree	P
5	Surface boulders	P
6	Limestone pavement	P
7	Peat hags	P
10	Soil erosion	P
17	10-50% peat	
31	Cliff > 30m high	P
32	Cliff 5-30m high	P
34	Rocky/boulder shore	P
35	Pebble/gravel shore	P
36	Sandy shore (or dune)	P
37	Bare mud	P
38	Sea	P
51	Lake - natural	P
52	Lake - artificial	P
53	River	P
54	Canalised river	P
55	Canal	P
56	Stream	P
58	Other ditch	P
61	Signs of drainage	P
64	Levee	P
65	Bank < 1m	
66	Bank 1-5m	
67	Bank > 5m	
90	Burnt heather	
91	Regenerating heather	
92	Vigorous heather	
93	Heather mosaic	
94	Heather dominant	
95	Collapsing heather	
96	Mat heather	
97	Bushy heather	
98	Mop heather	
99	Dead heather	
101	Lowland agricultural grass	P
102	Upland grass	P
103	Moorland - grass	P
104	Moorland - shrub heath	P
105	Calcareous grassland	P
106	Maritime vegetation	P
107	Lowland heath	P
108	Aquatic macrophytes	P
110	Raised bog	P
111	Blanket bog	P
112	Valley bog	P

CODE	DESCRIPTION	
113	Fen	P
114	Marsh	P
115	Flush	P
116	Saltmarsh	P
117	Wheat	P
118	Barley	P
119	Oats	P
120	Sugar beat	P
121	Turnips/swedes/roots	P
122	Kale	P
123	Potatoes	P
124	Field beans	P
125	Peas	P
126	Maize	P
127	Rye	P
128	Oilseed rape	P
129	Other crop	P
130	Flowers	P
131	Commercial horticulture	P
132	Orchard	P
133	Unmanaged grass	P
134	Tall herb vegetation	P
136	Ley	
137	Unimproved grass	
138	Forbs 10-25% (grass)	
139	Forbs 25-50% (grass)	
140	Forbs > 50% (grass)	
141	Neglected	
142	Abandoned	
143	Ploughed	P
144	Burnt	
145	Mown	
146	Lolium multiflorum	
147	Lolium perenne	
148	Trifolium repens	
149	Dactylis glomerata	
150	Anthoxanthum odoratum	
151	Phleum pratense	
152	Cynosurus cristatus	
153	Holcus lanatus	
154	Agrostis tenuis	
155	Festuca ovina	
156	Pteridium aquilinum - dense	P
157	Pteridium aquilinum -	
158	Juncus effusus	
159	Deschampsiaflexuosa	
160	Nardus stricta	
161	Calluna vulgaris	
162	Vaccinium myrtilus	

CODE	DESCRIPTION	
163	Molinia caerulea	
164	Eriophorum angustifolium	
165	Eriophorum vaginatum	
166	Trichorophorum caespitosum	
167	Sphagnum spp.	
168	Juncus squarrosus	
169	Juncus squarrosus	
170	Erica cinerea	
175	25-50%	
176	50-75%	
177	75-95%	
178	95-100%	
179	< 10cm	
180	10-30cm	
181	30-50cm	
182	0.5-1m	
183	1-1.5m	
184	> 1.5m	
185	Beef	
186	Dairy	
187	Breeders	
188	Dual purpose	
189	Sheep	
190	Goats	
191	Horses	
192	Pigs	
193	Silage	
194	Hay	
195	Deer	
196	Grouse	
197	No apparent use	
201	Individual trees	P
202	Scattered trees	P
203	Line of trees	P
204	Belt of trees	P
205	Clump of trees	P
206	Woodland/forest	P
207	Individual scrub species	P
208	Scattered scrub	P
209	Line of scrub	P
210	Patch of scrub	P
215	Closed canopy	
216	Canopies not touching	
218	Parkland	
221	Fir - Douglas	
222	Larch	
223	Pine - Corsican	
224	Pine - Lodgepole	
225	Pine – Scots	

CODE	DESCRIPTION	
226	Spruce - Norway	
227	Spruce - Sitka	
228	Unspecified conifer	
231	Alder	
232	Ash	
233	Beech	
234	Birch	
235	Bramble	
236	Elder	
237	Elm	
239	Gorse	
240	Hawthorn	
241	Hornbeam	
242	Lime	
243	Oak	
244	Poplar	
245	Rowan	
246	Sweet chestnut	
247	Sycamore	
248	Willow	
250	Mixed broadleaf	
251	Mixed conifer	
252	Unspecified broadleaf	
256	25-50%	
257	50-75%	
258	75-95%	
259	95-100%	
261	1-4 years	
262	5-20 years	
263	20-100 years	
264	> 100 years	
266	Timber production	
267	Landscape	
268	Sporting/game	
269	Public recreation	
270	Nature conservation	
271	Shelter	
275	Well managed	
276	Unmanaged - thriving	
277	Unmanaged - improvable	
278	Declining	
281	Felling/stumps	
282	Natural regeneration	
283	Underplanting	
284	Planted	
286	Staked trees	
287	Tree protectors	
288	Fenced (single trees)	
289	Windblow	
290	Dead standing trees	P

CODE	DESCRIPTION	
291	Regrowth - cut stump	
292	Grazing (stock)	
293	Ride/firebreak	P
294	Bracken - dense	
295	Bracken - scattered	
322	Hedge > 50% other species	P
331	Stone bank	P
341	> 2m high	
358	Uncut	
364	Bracken present	
401	Building	P
402	Garden/grounds with trees	P
403	Garden/grounds without trees	P
404	Public open space	P
405	Amenity grass > 1ha	P
406	Allotments	P
407	Car park	P
409	Graden centre/nursery	P
410	Embankment	P
411	Other land	P
421	Residential	
422	Commercial	
423	Industrial	
424	Public service & facilities	
425	Institutional	
426	Educational/cultural	
427	Religious	
428	Agricultural	
429	Sporting/recreational	
430	Waste - domestic	P
431	Waste - industrial	P
432	Quarry/mine	
433	Gravel pit	
441	New	
442	Vacant	
443	Derelict	
451	Railway track/land	P
452	Road (tarmac)	P
453	Verge < 1m	
454	Verge 1-5m	
455	Verge > 5m	
456	Constructed track	P
457	Unconstructed track	P
458	Footpath (exclusive)	P
459	Footpath (other)	P
460	Satisfactory throughout	
501	School playing fields	P
502	Other playing fields	P
503	Golf course	P
507	Static caravan(s)	P

CODE	DESCRIPTION	
508	Touring caravan park	P
511	Other designated area	P
521	Horsiculture	P
522	Angling	P
524	Other recreation	P
601	Dry heath	P
602	Wet heath	P
605	Rock crevice veg (maritime)	
608	Maritime grassland	
609	Maritime heath	
610	Strandline vegetation	
613	Stable yellow dune	
615	Dune slacks	
616	Inundation grassland (maritime)	
621	Pioneer saltmarsh	
622	Low marsh	
623	Middle/upper marsh	
625	Weeds>10%	
626	Weeds>25%	
627	Weeds>50%	
628	Festuca rubra	
630	Cattle	
631	Setaside	
640	Blackthorn	
641	Hazel	
645	Glade	P
651	10-25% (1ry codes only)	
652	26-50%	
653	51-75%	
654	76-95%	
655	96-100%	
678	Calcareous	
777	Unknown	P
888	New to map	P

** P = Primary Code*

Note: Surveyors were allowed to add unique codes in each Square, therefore certain codes will not be listed here

Appendix III - Boundary Codes

CODE	DESCRIPTION		CODE	DESCRIPTION		CODE	DESCRIPTION	
34	Rocky/boulder shore	P	237	Elm		352	Not stockproof	
38	Sea	P	238	Field maple		353	Filled gaps < 10%	
52	Lake - artificial	P	239	Gorse		354	Filled gaps > 10%	
53	River	P	240	Hawthorn		355	Signs of replacement	
55	Canal	P	241	Hornbeam		356	Signs of removal	
56	Stream	P	242	Lime		357	Trimmed	
57	Roadside ditch	P	243	Oak		358	Uncut	
58	Other ditch	P	244	Poplar		359	Derelict	
62	Waterfall	P	245	Rowan		360	Line of relict hedge	
63	Gorge	P	246	Sweet chestnut		361	Laying	
65	Bank < 1m		247	Sycamore		362	Flailing	
66	Bank 1-5m		248	Willow		363	Regrowth from stumps	
67	Bank > 5m		250	Mixed broadleaf		364	Bracken present	P
94	Heather dominant		251	Mixed conifer		410	Embankment	
112	Valley bog	P	252	Unspecified broadleaf		421	Residential	
113	Fen	P	256	25-50%		443	Derelict	P
134	Tall herb vegetation	P	257	50-75%		457	Unconstructed track	
149	Dactylis glomerata		258	75-95%		640	Blackthorn	
151	Phleum pratense		259	95-100%		641	Hazel	
162	Vaccinium myrtilus		261	1-4 years		651	10-25% (1ry codes only)	
175	25-50%		262	5-20 years		652	26-50%	
176	50-75%		263	20-100 years		653	51-75%	
178	95-100%		264	> 100 years		654	76-95%	
180	10-30cm		267	Landscape		655	96-100%	
181	30-50cm		271	Shelter		888	New to map	P
182	0.5-1m		282	Natural regeneration		999	No longer on map	
183	1-1.5m		284	Planted				
184	> 1.5m		286	Staked trees				
189	Sheep		287	Tree protectors				
201	Individual trees	P	290	Dead standing trees	P			
203	Line of trees	P	291	Regrowth - cut stump				
204	Belt of trees	P	295	Bracken - scattered				
205	Clump of trees	P	301	Dry-stone wall	P			
207	Individual scrub species	P	302	Mortared wall	P			
208	Scattered scrub	P	303	Other wall	P			
209	Line of scrub	P	311	Fence - wood only	P			
210	Patch of scrub	P	312	Fence - iron only	P			
215	Closed canopy		313	Fence - wire on posts	P			
216	Canopies not touching		314	Other fence	P			
217	Hedgerow		321	Hedge > 50% hawthorn	P			
222	Larch		322	Hedge > 50% other	P			
225	Pine - Scots		323	Mixed hedge	P			
228	Unspecified conifer		331	Stone bank	P			
231	Alder		332	Earth bank	P			
232	Ash		333	Grass strip only	P			
233	Beech		341	> 2m high				
234	Birch		342	1-2m high				
235	Bramble		343	< 1m high				
236	Elder		351	Stockproof				

* P = Primary Code

Note: Surveyors were allowed to add unique codes in each Square, therefore certain codes will not be listed here

Appendix IV: LUSE Codes

WLF	**Woody linear feature (hedge)**
AN	**Agriculture/Natural Vegetation**
CF	**Coastal Feature**
FO	**Forestry**
FOU	**Forestry Use**
IL	**Inland Physiography**
IW	**Inland Water**
RE	**Recreation**
ST	**Structures**
TR	**Transport**
AC	**Agricultural Crops**
FOF	**Forestry Feature**
GS	**Grass Strip**
B	**Bank**
F	**Fence**
W	**Wall**
US	**Un-surveyed/Missing Data**
AGU	**Agriculture/Natural Vegetation Use**

Appendix v: EIDC Guidance for supporting metadata

Guidance as provided by EIDC via https://eidc.ac.uk/deposit/supportingDocumentation

Supporting documentation is essential as it helps others understand a dataset and supports its potential re-use. This guidance document outlines what supporting information is required when submitting a dataset for deposit with the EIDC.

What information is required?

The information required for supporting documentation is likely to already exist. For example, in technical reports, in a data management plan, on a project website or wiki, or the information may be integral to the data itself (e.g. netCDF format).

Supporting documentation should provide information on the following areas:

A brief description/overview of the data being described

This is a summary that allows the reader to determine the relevance and usefulness of the data. The text should be concise but should contain sufficient detail to allow the reader to ascertain rapidly the scope and limitations of the resource.

To help organise thinking, you may like to use the following structure:

What has been recorded and what form does the data take? This should immediately convey to the reader precisely what the data is.
Where were the data collected? This should include (where appropriate) whether the coverage is gridded or scattered data; whether the coverage is even or very variable
When were the data were collected?
How were the data collected? A brief description of methods and instrumentation used.
Why were the data collected? For what purpose?
Who was responsible for the collection and interpretation of data?
Completeness. Are any data absent from the dataset? Explain which data are included or excluded and why.

Experimental design/sampling regime

Metadata should be provided which details the experimental design and/or sampling regime, where applicable. This should include information on:

- The feature(s) of interest - including feature type, feature name, relevant geographical information and grid or reference system used e.g. location, aspect, elevation, surface area, volume, etc.
- The treatments applied - including details of how treatments were applied/created/managed or verified, where relevant.
- Replication - details of any sample/observation replication, including explanations for any missing samples/observations.
- Controls - information on any control methods employed.
- The periodicity of sampling - details of the time period covered by the dataset, date and frequency of sample collection/observation recording and any reasons for missed sampling/observations.
- The number of samples/observations - the total overall number of samples/observations collected, if not already included as part of metadata relating to treatments and replication.
- Collection/Generation/Transformation Methods

- Information should be provided covering methods used for collection of samples or observations. This should include details of sampling/observation locations
- Techniques employed for physical collection of samples or measurement of observations
- Sample storage/treatment or recording of observations
- Standard Operating Procedures (SOPs) for specific techniques and/or references for methods used, if available
- Date of analysis of samples if different from date of sampling
- Alternatively, where data values are derived/generated/transformed, then details of how this was achieved should be provided. For model output, this should include information relating to the key points of the theory forming the basis of the model. The type of model used (e.g. ordinary differential equations, partial differential equations, compartment model) should be documented and any relevant technical information (e.g. operating system(s) and programming language) and/or mathematical information (e.g. input and output) used to generate the output also documented.

Fieldwork and/or laboratory instrumentation
Information should be supplied on instruments/machines used for collection/analysis of samples/observations where relevant. This should include the type, make, model and serial number of each particular instrument/machine, where known.

Calibration steps and values
Details of the steps taken to calibrate any instruments/machines used, including use of any blanks, and the values used for calibration should be provided.

Nature and units of recorded values
Information should be provided describing the nature of the recorded values contained and the units used sufficient to unambiguously define what has been measured and recorded in the dataset. Details should include description of the parameters, determinands, variables, valid range of values, lowest level of detection, units etc.

Analytical Methods
Full descriptions of any analytical methods used to generate the data values contained in the dataset should be included. These should detail any reagents and the specific conditions required for each analysis, and provide sufficient detail to enable replication of the methods used for analysis if desired.

Quality control
Any quality control measures undertaken to ensure the quality of the data values included in the dataset should be detailed e.g. methods of quality control, explanation of quality codes, factors affecting the data, etc.

Details of data structure
Details of the structure of the dataset should also be provided, covering the order in which variables appear within the dataset. For example:

This dataset comprises six csv files entitled xxx, yyy, zzz etc. The first csv file has five columns labelled aa, bb etc.

Miscellaneous
Any additional information necessary to expand on that given in the discovery metadata record.

How should I supply the information?
In order to meet funders' expectations and the needs of future research, supporting information must be openly accessible in perpetuity.

In order to guarantee this the EIDC requires supporting documentation to be submitted at the same time as submission of data. Unfortunately it is not acceptable to link to pages/documents on non-EIDC websites, or to include documents with hyperlinks to external websites. This is because we are unable to guarantee that those websites/pages will exist in perpetuity.

Format
Our preferred format is rich-text format (rtf). However we will also accept pdfs, html or plain text files.
If your information is stored in a proprietary format such as Microsoft Word it should be converted to a preferred format.

Stand-alone documents
Supporting documentation should be supplied separately from data. This is to ensure detailed metadata are available without the need to download the data, thus permitting users to make an informed decision about the utility of data prior to data access, and in perpetuity. Making the data and metadata available separately ensures the EIDC are able to securely store the data and ensure it remains unchanged whilst being able to improve the quality and usefulness of the contextual metadata, as required.

If the information is integral to the data itself (e.g. netCDF) it should be copied into a document in a preferred format.

Filenames
Names of files supplied must not include any spaces or non-standard characters.

Additional notes
More information is available in the NERC Data Policy guidance notes.

Depositors often ask if they can provide their scientific papers published through journals as supporting documentation for the dataset. Unfortunately these do not meet our requirements for a number of reasons:

- Copyright and distribution restrictions - there may be rights issues preventing us from making an article available publicly
- Lacking dataset description - the dataset is not described adequately in a publication designed to describe a scientific research outcome
- Access cannot be guaranteed in perpetuity - the EIDC cannot guarantee permanent, open access to external websites over which we have no control.

However, research papers may contain a lot of useful, pertinent information. In which case, we recommend copying the relevant content into a stand-alone document which can be supplied to us.

Appendix vi: Full list of my publications to date

http://nora.nerc.ac.uk/view/author/2811.html

*Papers submitted for consideration for PhD by Published Work are marked *.*

Publication – Article

Baxendale, C. L., Ostle, N. J., **Wood, C. M.**, Oakley, S., & Ward, S. E. (2016). Can digital image classification be used as a standardised method for surveying peatland vegetation cover? *Ecological Indicators, 68*, 150-156.

*Bunce, R. G. H., **Wood, C. M.**, & Smart, S. M. (2018a). The ecology of British upland landscapes. I. Composition of landscapes, habitats, vegetation and species. *Journal of Landscape Ecology, 11*(3), 120-139.

Bunce, R. G. H., **Wood, C. M.**, & Smart, S. M. (2018b). The ecology of British upland landscapes. II. The influence of policy on the current character of the uplands and the potential for change. *Journal of Landscape Ecology, 11*(3), 140-154.

Bunce, R. G. H., **Wood, C. M.**, Smart, S. M., Oakley, R., Browning, G., Daniels, M. J., Ashmole, P., Cresswell, J., & Holl, K. (2014). The landscape ecological impact of afforestation on the British uplands and some initiatives to restore native woodland cover. *Journal of Landscape Ecology, 7*(2), 5-24.

Maskell, L., Smart, S., Norton, L., & **Wood, C.** (2012). Vegetation database of Great Britain: Countryside Survey. *Biodiversity & Ecology, 4*.

McNamara, N. P., Plant, T., Oakley, S., Ward, S., **Wood, C.**, & Ostle, N. (2008). Gully hotspot contribution to landscape methane (CH4) and carbon dioxide (CO2) fluxes in a northern peatland. *Science of the Total Environment, 404*(2-3), 354-360.

Norton, L. R., Maskell, L. C., Smart, S. S., Dunbar, M. J., Emmett, B. A., Carey, P. D., Williams, P., Crowe, A., Chandler, K., Scott, W. A., & **Wood, C. M.** (2012). Measuring stock and change in the GB countryside for policy: key findings and developments from the Countryside Survey 2007 field survey. *Journal of Environmental Management, 113*, 117-127.

Norton, L. R., Smart, S. M., Maskell, L. C., Henrys, P. A., **Wood, C. M.**, Keith, A. M., Emmett, B. A., Cosby, B. J., Thomas, A., Scholefield, P. A., Greene, S., Morton, R. D., & Rowland, C. S. (2018). Identifying effective approaches for monitoring national natural capital for policy use. *Ecosystem Services, 30*(A), 98-106.

Reynolds, B., Chamberlain, P. M., Poskitt, J., **Wood, C. M.**, Scott, W. A., Rowe, E. C., Robinson, D. A., Frogbrook, Z. L., Keith, A. M., Henrys, P. A., Black, H. I. J., & Emmett, B. A. (2013). Countryside Survey: national "soil change" 1978-2007 for topsoils in Great Britain - acidity, carbon, and total nitrogen status. *Vadose Zone Journal, 12*(2).

Ritz, K., Black, H. I. J., Campbell, C. D., Harris, J. A., & **Wood, C.** (2009). Selecting biological indicators for monitoring soils: a framework for balancing scientific opinion to assist policy development. *Ecological Indicators, 9*(6), 1212-1221.

Rose, R., Monteith, D. T., Henrys, P., Smart, S., **Wood, C.**, Morecroft, M., Andrews, C., Beaumont, D., Benham, S., Bowmaker, V., Corbett, S., Dick, J., Dodd, B., Dodd, N., Flexen, M., McKenna, C., McMillan, S., Pallett, D., Rennie, S., Schafer, S., Scott, T., Sherrin, L., Turner, A., & Watson, H. (2016). Evidence for increases in vegetation species richness across UK Environmental Change Network sites linked to changes in air pollution and weather patterns. *Ecological Indicators, 68*, 52-62.

Rose, R. J., Roadknight, C. M., Barber, M. L., **Wood, C. M.**, & Marshall, I. W. (2009). Quantifying grazer activity in remote upland areas using digital cameras and automated image analysis. *Looking to the Hills: Newsletter of the Uplands Lead Co-ordination Network*(16), 6-10.

Scholefield, P., Morton, D., McShane, G., Carrasco, L., Whitfield, M. G., Rowland, C., Rose, R., **Wood, C.**, Tebbs, E., Dodd, B., & Monteith, D. (2019). Estimating habitat extent and carbon loss from an eroded northern blanket bog using UAV derived imagery and topography. *Progress in Physical Geography, 43*(2), 282-298.

Smart, S. M., Henrys, P. A., Norton, L. R., Wallace, H., **Wood, C. M.**, Williams, B., & Bunce, R. G. H. (2017). Changes in the frequency of common plant species across linear features in Wales from 1990 to 2016: implications for potential delivery of ecosystem services. *New Journal of Botany, 7*(2-3), 112-124.

Spake, R., Bellamy, C., Graham, L. J., Watts, K., Wilson, T., Norton, L. R., **Wood, C. M.**, Schmucki, R., Bullock, J. M., & Eigenbrod, F. (2019). An analytical framework for spatially targeted management of natural capital. *Nature Sustainability, 2*(2), 90-97.

***Wood, C. M.**, & Bunce, R. G. H. (2016a). Ecological survey of the native pinewoods of Scotland 1971. *Earth System Science Data, 8*(1), 177-189.

***Wood, C. M.**, & Bunce, R. G. H. (2016b). Survey of the terrestrial habitats and vegetation of Shetland, 1974 – a framework for long-term ecological monitoring. *Earth System Science Data, 8*(1), 89-103.

***Wood, C. M.**, Bunce, R. G. H., Norton, L. R., Maskell, L. C., Smart, S. M., Scott, W. A., Henrys, P. A., Howard, D. C., Wright, S. M., Brown, M. J., Scott, R. J., Stuart, R. C., & Watkins, J. W. (2018a). Ecological landscape elements: long-term monitoring in Great Britain, the Countryside Survey 1978-2007 and beyond. *Earth System Science Data, 10*(2), 745-763.

***Wood, C. M.**, Bunce, R. G. H., Norton, L. R., Smart, S. M., & Barr, C. J. (2018b). Land cover and vegetation data from an ecological survey of `key habitat' landscapes in England, 1992-1993. *Earth System Science Data, 10*(2), 899-918.

*Wood, C. M., Smart, S. M., & Bunce, R. G. H. (2015). Woodland Survey of Great Britain 1971-2001. *Earth System Science Data, 7*(2), 203-214.
*Wood, C. M., Smart, S. M., Bunce, R. G. H., Norton, L. R., Maskell, L. C., Howard, D. C., Scott, W. A., & Henrys, P. A. (2017). Long-term vegetation monitoring in Great Britain - the Countryside Survey 1978-2007 and beyond. *Earth System Science Data, 9*(2), 445-459.

Publication - Report

Barr, C. J., & **Wood, C. M.** (2011). The Sampling Strategy for Countryside Survey (up to 2007). Revised and Updated from: 'The Sampling Strategy for Countryside Survey', C.J. Barr, September 1998. DETR CONTRACT No. CR0212.http://nora.nerc.ac.uk/id/eprint/19487/
Black, H. I. J., Ritz, K., Campbell, C. D., Harris, J. A., **Wood, C.**, Chamberlain, P. M., Parekh, N., Towers, W., & Scott, A. (2008). SQID: Prioritising biological indicators of soil quality for deployment in a national-scale soil monitoring scheme. Summary report.http://nora.nerc.ac.uk/id/eprint/8108/
Black, H. I. J., Ritz, K., Harris, J. A., Cameron, C. M., Campbell, C. D., Chamberlain, P. M., Creamer, R., Pawlett, M., **Wood, C.**, & Singh, B. K. (2011). Scoping biological indicators of soil quality Phase II. Defra Final Contract Report SP0534.http://nora.nerc.ac.uk/id/eprint/503307/
Emmett, B., Abdalla, M., Anthony, S., August, T., Barrett, G., Biggs, J., Botham, M., Bradley, D., Brown, M., Carter, H., Chadwick, D., Cigna, F., Cooper, D., Cooper, J., Cosby, J., Creer, S., Cross, P., Edwards, F., Edwards, M., Evans, C., Ewald, N., Fitton, A., Garbutt, A., Grebby, S., Greene, S., Halfpenney, I., Hall, J., Harrison, S., Harrower, C., Henrys, P., Hobson, R., Hughes, S., Isaac, N., Jackson, B., Jarvis, S., Jones, D., Keith, A., Kelly, M., Korenko, J., Lallias, D., Leaver, D., Lebron, I., Malcolm, H., Maskell, L., McDonald, J., Moxley, J., Norton, L., O'Hare, M., Owen, A., Dos Santos Pereira, G., Peyton, J., Powney, G., Pywell, R., Rawlins, B., Robinson, D., Rorke, S., Rowland, C., Roy, D., Scarlett, P., Scholefield, P., Scott, A., Scott, L., Scott, R., Siriwardena, G., Smart, S., Smith, P., Swetnam, R., Taylor, R., Tordoff, G., Van Breda, J., Vincent, H., Wagner, M., Waters, E., Watkins, J., White, J., Williams, B., **Wood, C.**, & Wright, S. (2014). Glastir Monitoring & Evaluation Programme. First year annual report. In B. Williams (Ed.). Bangor.http://nora.nerc.ac.uk/id/eprint/506219/
Emmett, B. A., Frogbrook, Z. L., Chamberlain, P. M., Griffiths, R., Pickup, R., Poskitt, J., Reynolds, B., Rowe, E., Rowland, P., Wilson, J., & **Wood, C. M.** (2008). Countryside Survey. Soils Manual.http://nora.nerc.ac.uk/id/eprint/5201/
Emmett, B. A., Reynolds, B., Chamberlain, P. M., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A. J., Poskitt, J., Potter, E., Robinson, D. A., Scott, A., **Wood, C.**, & Woods, C. (2010). Countryside Survey: Soils Report from 2007.http://nora.nerc.ac.uk/id/eprint/9354/
Emmett, B. E., Abdalla, M., Anthony, S., Astbury, S., August, T., Barrett, G., Beckmann, B., Biggs, J., Botham, M., Bradley, D., Brown, M., Burden, A., Carter, H., Chadwick, D., Cigna, F., Collier, R., Cooper, D., Cooper, J., Cosby, B. J., Creer, S., Cross, P., Dadam, D., Edwards, F., Edwards, M., Evans, C., Ewald, N., Fitton, A., Garbutt, A., Giampieri, C., Gooday, R., Grebby, S., Greene, S., Halfpenney, I., Hall, J., Harrison, S., Harrower, C., Henrys, P., Hobson, R., Hughes, P., Hughes, S., Illian, J., Isaac, N., Jackson, B., Jarvis, S., Jones, D. L., Jones, P., Keith, A., Kelly, M., Kneebone, N., Korenko, J., Lallias, D., Leaver, D., Robinson, I., Malcolm, H., Maskell, L., McDonald, J., Moxley, J., Norton, L., O'Hare, M., Oliver, T., Owen, A., Parkhill, K. A., Pereira, M., Peyton, J., Pogson, M., Powney, G., Pritchard, N., Prochorskaite, A., Prosser, M., Pywell, R., Rawlins, B., Reuland, O., Richards, M., Robinson, D. A., Rorke, S., Rowland, C., Roy, D., Scarlett, P., Scholefield, P., Scott, A., Scott, L., Scott, R., Sharps, K., Siriwardena, G., Smart, S., Smith, G., Smith, P., Stopps, J., Swetnam, R., Taft, H., Taylor, R., Tebbs, E., Thomas, A., Todd-Jones, C., Tordoff, G., Turner, G., Van Breda, J., Vincent, H., Wagner, M., Waters, E., Walker-Springett, K., Wallace, H., Watkins, J., Webb, G., White, J., Whitworth, E., Williams, B., Williams, P., **Wood, C.**, & Wright, S. (2017). Glastir Monitoring & Evaluation Programme. Final report. Bangor, UK.http://nora.nerc.ac.uk/id/eprint/518194/
Emmett, B. E., Abdalla, M., Anthony, S., Astbury, S., August, T., Barrett, G., Beckmann, B., Biggs, J., Botham, M., Bradley, D., Chadwick, D., Collier, R., Cooper, D., Cooper, J., Cosby, B. J., Creer, S., Cross, P., Dadam, D., Edwards, F., Edwards, M., Evans, C., Ewald, N., Garbutt, A., Giampieri, C., Goodwin, A., Grebby, S., Greene, S., Halfpenney, I., Hall, J., Harrower, C., Henrys, P., Hobson, R., Hughes, P., Hughes, S., Isaac, N., Jackson, B., Jarvis, S., Jones, D. L., Jones, P., Keith, A., Kelly, M., Kneebone, N., Lallias, D., Lebron, I., Malcolm, H., Maskell, L., MacDonald, J., Maxwell, D., Moxley, J., Norton, L., Oliver, T., Owen, A., Parkhill, K. A., Pereira, M. G., Peyton, J., Powney, G., Prochorskaite, A., Rawlins, B., Reuland, O., Robinson, D. A., Rorke, S., Rowland, C., Roy, D., Scarlett, P., Scholefield, P., Scott, L., Smith, G. R., Siriwardena, G., Smart, S., Smith, P., Swetnam, R., Taft, H., Taylor, R., Tebbs, E., Thomas, A., Tordoff, G., Turner, G., Van Breda, J., Vincent, H., Wagner, M., Waters, E., Walker-Springett, K., Wallace, H., Webb, G., Williams, B., Williams, P., & **Wood, C.** (2015). Glastir Monitoring & Evaluation Programme. Second year annual report. In B. Emmett & B. Williams (Eds.). Bangor.http://nora.nerc.ac.uk/id/eprint/512331/
Maskell, L. C., Norton, L. R., Smart, S. M., Carey, P. D., Murphy, J., Chamberlain, P. M., **Wood, C. M.**, Bunce, R. G. H., & Barr, C. J. (2008a). Countryside Survey. Field Mapping Handbook.http://nora.nerc.ac.uk/id/eprint/5194/
Maskell, L. C., Norton, L. R., Smart, S. M., Scott, R., Carey, P. D., Murphy, J., Chamberlain, P. M., **Wood, C. M.**, Barr, C. J., & Bunce, R. G. H. (2008b). Countryside Survey. Vegetation Plots Handbook.http://nora.nerc.ac.uk/id/eprint/5196/
Mayr, T., Black, H., Towers, W., Palmer, R., Cooke, H., Freeman, M., Hornung, M., **Wood, C. M.**, Wright, S., Lilly, A., DeGroote, J., Jones, M., Cranfield University, The Macaulay Institute, Ecology, C. f., & Hydrology. (2006). Novel

methods for spatial prediction of soil functions within landscapes (SP0531).http://nora.nerc.ac.uk/id/eprint/11787/
Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., Smith, G., Wadsworth, R., & Simpson, I. (2011). Final Report for LCM2007 - the new UK land cover map. Countryside Survey Technical Report No 11/07.http://nora.nerc.ac.uk/id/eprint/14854/
Smart, S. M., Allen, D., Murphy, J., Carey, P. D., Emmett, B. A., Reynolds, B., Simpson, I. C., Evans, R. A., Skates, J., Scott, W. A., Maskell, L. C., Norton, L. R., Rossall, M. J., & **Wood, C.** (2009). Countryside Survey: Wales results from 2007.http://nora.nerc.ac.uk/id/eprint/7832/
Smart, S. M., Kirby, K., Corney, P., & **Wood, C. M.** (2013). Woodlands Survey of Great Britain 1971-2001: dataset documentation.http://nora.nerc.ac.uk/id/eprint/504075/
Wood, C. (2006a). Countryside Survey 2007 (Soils) Preparatory Phase II: Soil bulk density sampling. CEH Lancaster.http://nora.nerc.ac.uk/id/eprint/503786/
Wood, C. (2006b). Database documentation. MASQ (Monitoring and Assessing Soil Quality) database of UK soil chemistry and biology data 1978-2000 (pp. 85pp).http://nora.nerc.ac.uk/id/eprint/4113/
Wood, C. (2009). Database Documentation. Countryside Survey (Soils Component). A database of UK Soil Chemistry and Biology data 1978 - 2007.http://nora.nerc.ac.uk/id/eprint/8744/
Wood, C., Smart, S., Ecology, C. f., & Hydrology. (2009). Database Documentation. Countryside Survey (Vegetation Plots Component) VegCIDS (CSVEG). A database of Countryside Vegetation data 1978-2007.http://nora.nerc.ac.uk/id/eprint/8732/
Wood, C. M. (2008). CS1978 – Data Rescue. Scoping study for digitizing Countryside Survey primary field data documents. Final Report.http://nora.nerc.ac.uk/id/eprint/8752/
Wood, C. M., & Adamson, J. K. (2014). Solute concentrations in water samples from clearfelled and standing Sitka spruce (Picea sitchensis) forest ecosystems, Kershope Forest: dataset documentation.http://nora.nerc.ac.uk/id/eprint/508256/
Wood, C. M., Caffrey, D., & Bunce, R. G. H. (2015a). Dataset documentation - Scottish Pinewoods Survey 1971 (Native Pinewood Survey).http://nora.nerc.ac.uk/id/eprint/511847/
Wood, C. M., Howard, D. C., Henrys, P. A., Bunce, R. G. H., & Smart, S. M. (2012). Countryside Survey: measuring habitat change over 30 years: 1978 data rescue - final report. Lancaster.http://nora.nerc.ac.uk/id/eprint/16880/
Wood, C. M., Norton, L. R., & Rowland, C. S. (2015b). What are the costs and benefits of using aerial photography to survey habitats in 1km squares? Lancaster, UK.http://nora.nerc.ac.uk/id/eprint/511330/

Publication - Conference Items

Black, H., Chamberlain, P., Cameron, C., Campbell, C., Creamer, R., Harris, J., Pawlett, M., Ritz, K., Singh, B., Thomas, N., Robinson, L., & **Wood, C.** (2008). *Can we establish useful biological indicators of soil quality for national-scale soil monitoring?* Paper presented at the EUROSOIL 2008, Vienna. http://nora.nerc.ac.uk/id/eprint/8743/
Bunce, R. G. H., Sepp, K., Kikas, T., & **Wood, C. M.** (2017). *A comparison between lowland arable landscapes in Great Britain with those in a Baltic State, Estonia*. Paper presented at the 25 Years of Landscape Ecology, Manchester. http://nora.nerc.ac.uk/id/eprint/517236/
Bunce, R. G. H., Smart, S. M., & **Wood, C. M.** (2012). *The impact of afforestation on the British uplands*. Paper presented at the 19th ialeUK Conference, Edinburgh. http://nora.nerc.ac.uk/id/eprint/500912/
Mayr, T., Black, H., Towers, W., Palmer, R., Cooke, H., Freeman, M., Hornung, M., **Wood, C. M.**, Wright, S., Lilly, A., DeGroote, J., & Jones, M. (2007). *Development of a soil function evaluation framework*. Paper presented at the Proceedings of BSSS Conference Soils - Fit for Purpose?, Aberystwyth. http://nora.nerc.ac.uk/id/eprint/8739/
Rowland, C., Morton, D., Marston, C., & **Wood, C.** (2011). *Land cover map 2007: using OBIA for LCM2007*. Paper presented at the Remote Sensing and Photogrammetry Society Annual Conference 2011 Earth Observation in a Changing World, Bournmouth University, UK. http://nora.nerc.ac.uk/id/eprint/16422/
Towers, W., Mayr, T., Black, H., Palmer, R., Cooke, H., Freeman, M., Hornung, M., **Wood, C. M.**, Wright, S., Lilly, A., DeGroote, J., & Jones, M. (2008). *What would we like our soils to do and how do we decide?* Paper presented at the Eurosoil 2008, Vienna. http://nora.nerc.ac.uk/id/eprint/8742/
Wood, C., & Bunce, R. G. H. (2017). *Characterising the upland landscapes of Great Britain*. Paper presented at the 25 Years of Landscape Ecology, Manchester. http://nora.nerc.ac.uk/id/eprint/517235/

Publication - Book Sections

Black, H. I. J., Chamberlain, P. M., Creamer, R., Campbell, C., Harris, J. A., Pawlett, M., Ritz, K., & **Wood, C. M.** (2010). The function of soil biodiversity as indicators of soil quality: Insights from the UK Defra SQID project. In R. Gilkes & N. Prakongkep (Eds.), *Soil solutions for a changing world* (pp. 108-110): International Union of Soil Sciences.
Keith, A. M., Griffiths, R. I., Henrys, P. A., Hughes, S., Lebron, I., Maskell, L. C., Ogle, S. M., Robinson, D. A., Rowe, E. C., Smart, S. M., Spurgeon, D., **Wood, C. M.**, & Emmett, B. A. (2015). Monitoring soil natural capital and ecosystem services by using large-scale survey data. In M. Stromberger, N. Comerford & D. Lindbo (Eds.), *Soil ecosystems services*. Madison, Wis.: Soil Science Society of America.

Rowland, C. S., Morton, R. D., Murfin, M., & **Wood, C. M.** (2013). Developing land cover and land cover change mapping methods for the UK: preliminary results. In L. Ouwehand (Ed.), *Proceedings of ESA Living Planet Symposium*. Noordwijk, The Netherlands: ESA Communications.

Output (Electronic) - Datasets

Adamson, J., Benefield, C., Elfyn Hughes, W., & **Wood, C.** (2017). Terrestrial habitat and vegetation data from the marginal uplands of Cumbria, 1978.http://nora.nerc.ac.uk/id/eprint/517799/

Barr, C. J., Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Clarke, R., Crowe, A., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2014 -a). Countryside Survey estimates of Broad Habitat area change in Great Britain between 1990,1998 and 2007.http://nora.nerc.ac.uk/id/eprint/509062/

Barr, C. J., Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Clarke, R., Crowe, A., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014a). Countryside Survey mapped estimates of Broad Habitat area change in Great Britain between 1990 and 2007.http://nora.nerc.ac.uk/id/eprint/509073/

Barr, C. J., Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Crowe, A., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2015a). Countryside Survey mapped estimates of Broad Habitat area change in Great Britain between 1998 and 2007.http://nora.nerc.ac.uk/id/eprint/510939/

Barr, C. J., Bunce, R. G. H., Carey, P., Chandler, K., Crowe, A., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2016a). Landscape area data 2007 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/512783/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2015b). Countryside Survey mapped estimates of Broad Habitat area change in Great Britain between 1990 and 1998.http://nora.nerc.ac.uk/id/eprint/510942/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014b). Countryside Survey 1990 estimates of linear feature lengths in Great Britain.http://nora.nerc.ac.uk/id/eprint/508792/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014c). Countryside Survey 1990 mapped estimates of linear feature lengths in Great Britain.http://nora.nerc.ac.uk/id/eprint/508794/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2016b). Landscape area data 1990 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/512780/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2016c). Landscape point feature data 1990 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513577/

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D. C., Maskell, L. C., Ness, M. J., Norton, L. R., Scott, R. J., Scott, W. A., Smart, S. M., Stuart, R. C., **Wood, C. M**., and Wright, S. M.: Landscape point feature data 1990 [Countryside Survey], http://nora.nerc.ac.uk/id/eprint/513366/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2014 -b). Countryside Survey 1998 estimates of linear feature lengths in Great Britain.http://nora.nerc.ac.uk/id/eprint/508926/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2014 -c). Countryside Survey 1998 mapped estimates of linear feature lengths in Great Britain.http://nora.nerc.ac.uk/id/eprint/508973/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2016d). Landscape point feature data 1998 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513576/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2016 -b). Landscape area data 1998 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/512782/

Barr, C. J., Bunce, R. G. H., Clarke, R., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2016 -c). Landscape linear feature data 1998 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513367/

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014d). Countryside Survey 1990 estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/504658/

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014e). Countryside Survey 1990 mapped estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/504659/

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2014f). Countryside Survey 1998 estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/504664/

Barr, C. J., Bunce, R. G. H., Clarke, R. T., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Watkins, J., **Wood, C.**, & Wright, S. (2014g). Countryside Survey 1998 mapped estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/504663/

Barr, C. J., Bunce, R. G. H., Cummins, R. P., Hallam, C. J., Hornung, M., & **Wood, C.** (2017). Habitat and vegetation data from an ecological survey of terrestrial key habitats in England, 1992-1993.http://nora.nerc.ac.uk/id/eprint/517658/

Barr, C. J., Bunce, R. G. H., Gillespie, M. K., Hallam, C. J., Howard, D., Maskell, L., Ness, M. J., Norton, L., Scott, R., Stuart, R., Smart, S., & **Wood, C.** (2014h). Countryside Survey 1990 vegetation plot data.http://nora.nerc.ac.uk/id/eprint/505300/

Barr, C. J., Bunce, R. G. H., Gillespie, M. K., Howard, D., Maskell, L., Norton, L., Scott, R., Shield, E. R., Smart, S., Stuart, R., Watkins, J., & **Wood, C.** (2014i). Countryside Survey 1998 vegetation plot data.http://nora.nerc.ac.uk/id/eprint/505299/

Barr, C. J., Bunce, R. G. H., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Whittaker, H. A., **Wood, C.**, & Wright, S. (2014j). Countryside Survey 1984 estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/504657/

Barr, C. J., Bunce, R. G. H., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Whittaker, H. A., **Wood, C.**, & Wright, S. (2014k). Countryside Survey 1984 estimates of linear feature lengths in Great Britain.http://nora.nerc.ac.uk/id/eprint/509029/

Barr, C. J., Bunce, R. G. H., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Whittaker, H. A., **Wood, C.**, & Wright, S. (2014 -d). Countryside Survey 1984 mapped estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/504656/

Barr, C. J., Bunce, R. G. H., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Whittaker, H. A., **Wood, C.**, & Wright, S. (2016e). Landscape area data 1984 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/512779/

Barr, C. J., Bunce, R. G. H., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Whittaker, H. A., **Wood, C.**, & Wright, S. (2016f). Landscape linear feature data 1984 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513364/

Barr, C. J., Bunce, R. G. H., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Whittaker, H. A., **Wood, C.**, & Wright, S. (2016g). Landscape point feature data 1984 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513578/

Barr, C. J., Bunce, R. G. H., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., Whittaker, H. A., **Wood, C.**, & Wright, S. (2020). Countryside Survey 1984 mapped estimates of linear feature lengths in Great Britain.http://nora.nerc.ac.uk/id/eprint/509028/

Black, H., Chaplow, J., Ainsworth, G., Coward, P., Creamer, R., Thacker, S., Horne, J., Hornung, M., Kennedy, V. H., Raine, L., Osborn, D., Parekh, N., Parrington, J., Poskitt, J., Potter, E., Reeves, N., Rowland, P., Self, P., Turner, S., Watkins, J., **Wood, C.**, Woods, C., & Wright, J. (2016a). Soil metals data 1998 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513899/

Black, H., Chaplow, J., Ainsworth, G., Coward, P., Creamer, R., Thacker, S., Horne, J., Hornung, M., Kennedy, V. H., Raine, L., Osborn, D., Parekh, N., Parrington, J., Poskitt, J., Potter, E., Reeves, N., Rowland, P., Self, P., Turner, S., Watkins, J., **Wood, C.**, Woods, C., & Wright, J. (2020). Soil physico-chemical properties 1998 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513891/

Black, H., Chaplow, J., Creamer, R., Hornung, M., Monson, F., Poskitt, J., Potter, E., Self, P., Watkins, J., & **Wood, C.** (2016b). Soil invertebrate data 1998 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513889/

Botham, M., Peyton, J., Tordoff, G., Edwards, M., Alison, J., Astbury, S., Burden, A., Emmett, B., Garbutt, A., Henrys, P., Owen, A., Smart, S., Sharps, K., Williams, B., & **Wood, C.** (2020). Insect pollinator and flower surveys from the Glastir Monitoring and Evaluation Programme, Wales, 2013 - 2016.http://nora.nerc.ac.uk/id/eprint/527920/

Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Crowe, A., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014a). Countryside Survey 2007 estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/504823/

Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Crowe, A., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014b). Countryside Survey 2007 estimates of linear feature lengths in Great Britain.http://nora.nerc.ac.uk/id/eprint/508499/

Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Crowe, A., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014c). Countryside Survey 2007 mapped estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/504826/

Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Crowe, A., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2014). Countryside Survey 2007 mapped estimates of linear feature lengths in Great Britain.http://nora.nerc.ac.uk/id/eprint/508500/

Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Crowe, A., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2016a). Landscape linear feature data 2007 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513369/

Brown, M., Bunce, R. G. H., Carey, P., Chandler, K., Crowe, A., Maskell, L., Norton, L., Scott, R., Scott, A., Smart, S., Stuart, R., **Wood, C.**, & Wright, S. (2016b). Landscape point feature data 2007 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513365/

Bunce, R. G. H., Barr, C. J., Shaw, M. W., Bunce, F., Bassett, P., Quarmby, S., Robertson, S. M. C., Bowen, W. J. A., Morris, D., Millar, A., Brown, A. H. F., Horrill, A. D., Dodd, B., & **Wood, C.** (2017a). Ecological survey of the Spey Valley birch woods, 1973.http://nora.nerc.ac.uk/id/eprint/516417/

Bunce, R. G. H., Barr, C. J., **Wood, C.**, & Dodd, B. (2017). Habitat, vegetation, soil, tree and regeneration data from native pinewoods in Scotland, 1973.http://nora.nerc.ac.uk/id/eprint/517133/

Bunce, R. G. H., Bassett, P., & **Wood, C.** (2015a). Terrestrial habitat, vegetation and soil data from Shetland, 1974.http://nora.nerc.ac.uk/id/eprint/511975/

Bunce, R. G. H., Brown, A. H. F., Shaw, M. W., White, E. J., Gardiner, A., Procter, J. E. A., Birkett, J. S., McAllister, J., Smith, C., Parrington, J., Wrigley, J., & **Wood, C.** (2018). Tree, vegetation and soil information from an ecological survey of semi-natural woodlands in the English Lake District, 1969.http://nora.nerc.ac.uk/id/eprint/520169/

Bunce, R. G. H., Carey, P., Maskell, L., Norton, L., Scott, R., Smart, S., & **Wood, C.** (2014a). Countryside Survey 2007 vegetation plot data.http://nora.nerc.ac.uk/id/eprint/505297/

Bunce, R. G. H., Shaw, M. W., & **Wood, C.** (2015b). Habitat, vegetation, tree and soil data from Native Pinewoods in Scotland, 1971.http://nora.nerc.ac.uk/id/eprint/511858/

Bunce, R. G. H., Smith, R. S., Booy, M., & **Wood, C.** (2017b). Terrestrial habitat, vegetation and soil data from Cumbria, England, 1975.http://nora.nerc.ac.uk/id/eprint/517291/

Bunce, R. G. H., **Wood, C.**, Henrys, P., Smart, S., Howard, D., & Barr, C. J. (2014b). Countryside Survey 1978 estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/508029/

Bunce, R. G. H., **Wood, C.**, Henrys, P., Smart, S., Howard, D., & Barr, C. J. (2014c). Countryside Survey 1978 mapped estimates of Broad Habitat areas in Great Britain.http://nora.nerc.ac.uk/id/eprint/508027/

Bunce, R. G. H., **Wood, C.**, Henrys, P., Smart, S., Howard, D., & Barr, C. J. (2016). Landscape area data 1978 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/512778/

Carlisle, A., & **Wood, C.** (2014). Soil pH at Roudsea Wood National Nature Reserve 1961.http://nora.nerc.ac.uk/id/eprint/509128/

Donelly, R. J., English, N., **Wood, C.**, & Hebbert, E. (2019). Vegetation map of Roudsea Wood National Nature Reserve, 1962.http://nora.nerc.ac.uk/id/eprint/525359/

Emmett, B., Reynolds, B., Chamberlain, P., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Keith, A., Lawlor, A., Monson, F., Poskitt, J., Potter, E., Robinson, D., Scott, A., Thompson, N., Watts, R., **Wood, C.**, & Woods, C. (2016a). Soil invertebrate data 2007 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513883/

Emmett, B., Reynolds, B., Chamberlain, P., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A., Poskitt, J., Potter, E., Robinson, D., Scott, A., **Wood, C.**, & Woods, C. (2016b). Soil metals data 2007 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513882/

Emmett, B., Reynolds, B., Chamberlain, P., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A., Poskitt, J., Potter, E., Robinson, D., Scott, A., **Wood, C.**, & Woods, C. (2016c). Soil physico-chemical properties 2007 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/513887/

Emmett, B., Reynolds, B., Chamberlain, P., Rowe, E., Spurgeon, D., Brittain, S. A., Frogbrook, Z., Hughes, S., Lawlor, A., Poskitt, J., Potter, E., Robinson, D., Scott, A., **Wood, C.**, & Woods, C. (2016d). Topsoil mineralisable nitrogen (mineral-N) data 2007 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/514018/

Henrys, P., Keith, A., & **Wood, C.** (2016). Model estimates of aboveground carbon for Great Britain.http://nora.nerc.ac.uk/id/eprint/513441/

Keith, A., Astbury, S., Dodd, B., Emmett, B., Fitton, A., Garbutt, A., George, P., Marshall, R., Robinson, D., Williams, B., & **Wood, C.** (2019). Topsoil meso-fauna data from the Glastir Monitoring and Evaluation Programme, Wales 2013-2014.http://nora.nerc.ac.uk/id/eprint/525961/

Keith, A., Monson, F. D., Chaplow, J., Black, H., Self, P., Rollett, A., & **Wood, C.** (2018). Oribatid mites 1998 [Countryside Survey].http://nora.nerc.ac.uk/id/eprint/520106/

Maskell, L., Astbury, S., Burden, A., Emmett, B., Garbutt, A., Goodwin, A., Henrys, P., Jarvis, S., Norton, L., Owen, A., Sharps, K., Smart, S., Williams, B., **Wood, C.**, & Wright, S. (2020a). Landscape linear feature data from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/527238/

Maskell, L., Astbury, S., Burden, A., Emmett, B., Garbutt, A., Goodwin, A., Henrys, P., Jarvis, S., Norton, L., Owen, A., Sharps, K., Smart, S., Williams, B., **Wood, C.**, & Wright, S. (2020b). Landscape point feature data from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/527237/

Maskell, L., Astbury, S., Burden, A., Emmett, B., Garbutt, A., Goodwin, A., Henrys, P., Jarvis, S., Norton, L., Owen, A., Sharps, K., Smart, S., Williams, B., **Wood, C.**, & Wright, S. (2020). Landscape and habitat area data from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/527236/

Maskell, L., Henrys, P., Norton, L., Smart, S., & **Wood, C.** (2013a). Distribution of ash trees within areas less than half a hectare in Great Britain.http://nora.nerc.ac.uk/id/eprint/503962/

Maskell, L., Henrys, P., Norton, L., Smart, S., & **Wood, C.** (2013b). Distribution of individual ash trees in Great Britain.http://nora.nerc.ac.uk/id/eprint/503963/

Maskell, L., Henrys, P., Norton, L., Smart, S., & **Wood, C.** (2013). Distribution of ash trees in woody linear features in Great Britain.http://nora.nerc.ac.uk/id/eprint/503961/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., & Smith, G. (2011a). Land Cover Map 2007 (1km dominant aggregate class, GB) v1.2.http://nora.nerc.ac.uk/id/eprint/507962/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., & Smith, G. (2011b). Land Cover Map 2007 (1km dominant aggregate class, N.Ireland) v1.2.http://nora.nerc.ac.uk/id/eprint/507984/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., & Smith, G. (2011c). Land Cover Map 2007 (1km percentage aggregate class, GB) v1.2.http://nora.nerc.ac.uk/id/eprint/507961/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., & Smith, G. (2011d). Land Cover Map 2007 (1km percentage target class, GB) v1.2.http://nora.nerc.ac.uk/id/eprint/507960/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., & Smith, G. (2011e). Land Cover Map 2007 (25m raster, GB) v1.2.http://nora.nerc.ac.uk/id/eprint/507959/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., & Smith, G. (2011f). Land Cover Map 2007 (vector, GB) v1.2.http://nora.nerc.ac.uk/id/eprint/507963/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., Smith, G., Wadsworth, R., & Simpson, I. (2011g). Land Cover Map 2007 (1km dominant target class, N. Ireland).http://nora.nerc.ac.uk/id/eprint/507966/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., Smith, G., Wadsworth, R., & Simpson, I. (2011h). Land Cover Map 2007 (1km percentage target class, N.Ireland).http://nora.nerc.ac.uk/id/eprint/507967/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., Smith, G., Wadsworth, R., & Simpson, I. (2011i). Land Cover Map 2007 (1km raster, percentage aggregate class, N.Ireland).http://nora.nerc.ac.uk/id/eprint/507964/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., Smith, G., Wadsworth, R., & Simpson, I. (2011j). Land Cover Map 2007 (25m raster, NI).http://nora.nerc.ac.uk/id/eprint/507965/

Morton, D., Rowland, C., **Wood, C.**, Meek, L., Marston, C., Smith, G., Wadsworth, R., & Simpson, I. (2011k). Land Cover Map 2007 (Vector, N. Ireland).http://nora.nerc.ac.uk/id/eprint/507968/

Robinson, D., Alison, J., Andrews, C., Brentegani, M., Chetiu, N.-F., Dart, S., Emmett, B., Fitos, E., Garbutt, A., Gray, A., Henrys, P., Hunt, A., Keenan, P., Keith, A., Koblizek, E., Robinson, I., Millani Lopes Mazzetto, J., Pallett, D., Dos Santos Pereira, G., Pinder, A., Rose, R., Rowe, R., Scarlett, P., Seaton, F., Smart, S., Towill, J., Wagner, M., Williams, B., & **Wood, C.** (2020a). Topsoil physico-chemical properties from the UKCEH Countryside Survey, Great Britain, 2019.http://nora.nerc.ac.uk/id/eprint/528651/

Robinson, D., Astbury, S., Barrett, G., Burden, A., Carter, H., Emmett, B., Garbutt, A., Giampieri, C., Hall, J., Henrys, P., Hughes, S., Hunt, A., Jarvis, S., Jones, D., Keenan, P., Lebron, I., Nunez Pastrana, D., Owen, A., Patel, M., Dos Santos Pereira, G., Seaton, F., Sharps, K., Tanna, B., Thompson, N., Williams, B., & **Wood, C.** (2019). Topsoil physico-chemical properties from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/524171/

Robinson, I., Seaton, F., Barrett, G., Alison, J., Burden, A., Emmett, B., Garbutt, A., Robinson, D., Williams, B., & **Wood, C.** (2020b). Topsoil particle size distribution from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/527649/

Rose, R., **Wood, C.**, Turner, C., Whitfield, M., Ostle, N., Driscoll, B., Looker, A., & Looker, G. (2016). Vegetation survey of Moor House National Nature Reserve 2008-2009.http://nora.nerc.ac.uk/id/eprint/515074/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017a). Land Cover Map 2015 (1km dominant aggregate class, GB).http://nora.nerc.ac.uk/id/eprint/516848/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017b). Land Cover Map 2015 (1km dominant aggregate class, N. Ireland).http://nora.nerc.ac.uk/id/eprint/516845/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017c). Land Cover Map 2015 (1km dominant target class, GB).http://nora.nerc.ac.uk/id/eprint/516840/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017d). Land Cover Map 2015 (1km dominant target class, N. Ireland).http://nora.nerc.ac.uk/id/eprint/516846/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017e). Land Cover Map 2015 (1km percentage aggregate class, GB).http://nora.nerc.ac.uk/id/eprint/516851/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017f). Land Cover Map 2015 (1km percentage aggregate class, N. Ireland).http://nora.nerc.ac.uk/id/eprint/516842/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017g). Land Cover Map 2015 (1km percentage target class, GB).http://nora.nerc.ac.uk/id/eprint/516844/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017h). Land Cover Map 2015 (1km percentage target class, N. Ireland).http://nora.nerc.ac.uk/id/eprint/516847/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017i). Land Cover Map 2015 (25m raster, GB).http://nora.nerc.ac.uk/id/eprint/516839/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017j). Land Cover Map 2015 (25m raster, N. Ireland).http://nora.nerc.ac.uk/id/eprint/516849/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017k). Land Cover Map 2015 (vector, GB).http://nora.nerc.ac.uk/id/eprint/516841/

Rowland, C., Morton, D., Carrasco Tornero, L., McShane, G., O'Neil, A., & **Wood, C.** (2017l). Land Cover Map 2015 (vector, N. Ireland).http://nora.nerc.ac.uk/id/eprint/516850/

Scarlett, P., Edwards, F., Astbury, S., Burden, A., Carter, H., Emmett, B., Garbutt, A., Henrys, P., Keenan, P., Owen, A., Sharps, K., Williams, B., & **Wood, C.** (2020a). Headwater stream quality metrics from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/527989/

Scarlett, P., Edwards, F., Astbury, S., Burden, A., Carter, H., Emmett, B., Garbutt, A., Henrys, P., Keenan, P., Owen, A., Sharps, K., Williams, B., & **Wood, C.** (2020b). Pond quality metrics from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/527648/

Scott, R., Baker, P., Bell, D., Bennett, S., Boffey, C., Bohan, D. A., Brooks, D. R., Browne, E. L., Champion, G. T., Clark, S. J., Daniels, R. E., Dewar, A. J. G., Dewar, A. M., Fairfax, C., Firbank, L. G., Garner, B. H., Haughton, A. J., Hawes, C., Haylock, L. A., Heard, M., Hill, M., Horne, S. L., Hulmes, S., Mason, N. S., May, M. J., Norton, L., Nuttall, P., Osborne, J. L., Parish, A. M., Perry, J. N., Randle, Z., Rossall, M., Rothery, P., Roy, D., Sands, R. J. N., Singer, E. J., Skellern, M. P., Squire, G. R., Thomas, M. R., Walker, J. H., Walker, M. J., Woiwod, I. P., **Wood, C.**, & Young, M. W. (2012a). Farm scale evaluations of herbicide tolerant genetically modified crops - beet.http://nora.nerc.ac.uk/id/eprint/523094/

Scott, R., Baker, P., Bell, D., Bennett, S., Boffey, C., Bohan, D. A., Brooks, D. R., Browne, E. L., Champion, G. T., Clark, S. J., Daniels, R. E., Dewar, A. J. G., Dewar, A. M., Fairfax, C., Firbank, L. G., Garner, B. H., Haughton, A. J., Hawes, C., Haylock, L. A., Heard, M., Hill, M., Horne, S. L., Hulmes, S., Mason, N. S., May, M. J., Norton, L., Nuttall, P., Osborne, J. L., Parish, A. M., Perry, J. N., Randle, Z., Rossall, M., Rothery, P., Roy, D., Sands, R. J. N., Singer, E. J., Skellern, M. P., Squire, G. R., Thomas, M. R., Walker, J. H., Walker, M. J., Woiwod, I. P., **Wood, C.**, & Young, M. W. (2012b). Farm scale evaluations of herbicide tolerant genetically modified crops - maize.http://nora.nerc.ac.uk/id/eprint/523095/

Scott, R., Baker, P., Bell, D., Bennett, S., Boffey, C., Bohan, D. A., Brooks, D. R., Browne, E. L., Champion, G. T., Clark, S. J., Daniels, R. E., Dewar, A. J. G., Dewar, A. M., Fairfax, C., Firbank, L. G., Garner, B. H., Haughton, A. J., Hawes, C., Haylock, L. A., Heard, M., Hill, M., Horne, S. L., Hulmes, S., Mason, N. S., May, M. J., Norton, L., Nuttall, P., Osborne, J. L., Parish, A. M., Perry, J. N., Randle, Z., Rossall, M., Rothery, P., Roy, D., Sands, R. J. N., Singer, E. J., Skellern, M. P., Squire, G. R., Thomas, M. R., Walker, J. H., Walker, M. J., Woiwod, I. P., **Wood, C.**, & Young, M. W. (2012c). Farm scale evaluations of herbicide tolerant genetically modified crops - spring oilseed rape.http://nora.nerc.ac.uk/id/eprint/523096/

Scott, R., Baker, P., Bell, D., Bennett, S., Boffey, C., Bohan, D. A., Brooks, D. R., Browne, E. L., Champion, G. T., Clark, S. J., Daniels, R. E., Dewar, A. J. G., Dewar, A. M., Fairfax, C., Firbank, L. G., Garner, B. H., Haughton, A. J., Hawes, C., Haylock, L. A., Heard, M., Hill, M., Horne, S. L., Hulmes, S., Mason, N. S., May, M. J., Norton, L., Nuttall, P., Osborne, J. L., Parish, A. M., Perry, J. N., Randle, Z., Rossall, M., Rothery, P., Roy, D., Sands, R. J. N., Singer, E. J., Skellern, M. P., Squire, G. R., Thomas, M. R., Walker, J. H., Walker, M. J., Woiwod, I. P., **Wood, C.**, & Young, M. W. (2012d). Farm scale evaluations of herbicide tolerant genetically modified crops - winter oilseed rape.http://nora.nerc.ac.uk/id/eprint/523097/

Siriwardena, G. M., Taylor, R. M., Norfolk, D., Jones, K., Alison, J., Astbury, S., Burden, A., Emmett, B., Garbutt, A., Henrys, P., Owen, A., Smart, S., Sharps, K., Williams, B., & **Wood, C.** (2020). Bird counts from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/527919/

Smart, S., Astbury, S., Burden, A., Emmett, B., Garbutt, A., Henrys, P., Maskell, L., Owen, A., Sharps, K., Wallace, H., Williams, B., & **Wood, C.** (2020). Vegetation plot data from the Glastir Monitoring and Evaluation Programme, Wales 2013-2016.http://nora.nerc.ac.uk/id/eprint/526887/